

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

H. CRISTINA CHEN-OSTER, LISA PARISI, SHANNA ORLICH, ALLISON GAMBA, and MARY DE LUIS,

Plaintiffs,

-against-

GOLDMAN, SACHS & CO. and THE GOLDMAN SACHS GROUP, INC.,

Defendants.

ANALISA TORRES, District Judge:

Plaintiffs H. Cristina Chen-Oster, Shanna Orlich, Allison Gamba, and Mary De Luis,¹ representing a class of female employees of Goldman, Sachs & Co. and The Goldman Sachs Group, Inc. (collectively, “Goldman Sachs” or “Defendants”), filed this class action alleging gender discrimination under Title VII of the Civil Rights Act of 1964 (“Title VII”), 42 U.S.C. § 2000e *et seq.*, and the New York City Human Rights Law (the “NYCHRL”), N.Y.C. Admin. Code § 8-101 *et seq.*

Before the Court are the parties’ motions to exclude expert testimony, ECF Nos. 1184, 1187, Defendants’ motion for decertification of the class, ECF No. 1194, Defendants’ motion for summary judgment, ECF No. 1239, and Plaintiffs’ motion for partial summary judgment, ECF No. 1247. For the reasons stated below, the parties’ motions to exclude expert testimony are GRANTED in part, and DENIED in part; Defendants’ motion to decertify the class is DENIED; Plaintiffs’ motion for summary judgment is DENIED; and Defendants’ motion for summary judgment is GRANTED in part, and DENIED in part.

USDC SDNY
DOCUMENT
ELECTRONICALLY FILED
DOC #: _____
DATE FILED: 3/17/2022

10 Civ. 6950 (AT) (RWL)

ORDER

¹ Originally, Lisa Parisi was also a named plaintiff. Compl. ¶¶ 15–16, ECF No. 5. Because Parisi’s claims are subject to a binding arbitration agreement, see ECF No. 171; *Parisi v. Goldman, Sachs & Co.*, 710 F.3d 483 (2d Cir. 2013), she is no longer a plaintiff, see generally Second Am. Compl., ECF No. 411.

BACKGROUND²

I. Goldman Sachs

Goldman Sachs is a leading financial services company that has four revenue-generating divisions—Investment Banking, Investment Management, Securities, and Merchant Banking. Defs. 56.1 ¶¶ 1–6, ECF No. 1241. These divisions are split into numerous specialized business units. *See, e.g.*, ECF No. 265-1. Plaintiffs worked in three of the revenue generating divisions. *See* Class Cert. Order at 2–4, ECF No. 578. In these three divisions, Goldman Sachs employed two systems for evaluating employees, known as “360 review” and “quartiling.” Pls. 56.1 ¶ 1, ECF No. 1257. For promotions from Vice President to Managing Director, Goldman Sachs used a process called “cross-ruffing.” Class Cert. Order at 8–10.

A. Performance Evaluation

1. 360 Review

Goldman Sachs employees in the Investment Banking, Investment Management, and Securities Divisions underwent the 360 review process. Pls. 56.1 ¶ 1. First, each employee performed a self-evaluation and named 8 to 12 evaluators—subordinates, bosses, peers, and internal clients—with whom he or she had recently worked to provide reviews of his or her performance. *Id.* ¶ 2; Defs. 56.1 ¶ 13. Evaluators assessed employees on criteria that were the same across the three divisions in any given year. Pls. 56.1 ¶ 3. Finally, the employee’s manager received the reviews, added their own evaluations, created a narrative summary of all of the reviews, and discussed the overall review with the employee. Defs. 56.1 ¶¶ 18–19. In 2016, the 360 review process shifted to being primarily used as a development tool rather than a

² The following facts are drawn from the parties’ Rule 56.1 statements of undisputed facts, and the opposing party’s response. Disputed facts are so noted. Citations to a paragraph in the Rule 56.1 statement also include the opposing party’s response.

compensation determination tool. Landman Decl. ¶¶ 6, 20, ECF No. 1250-5; Landman Dep. at 82, 88, ECF No. 1289-8.

2. Quartiling

Goldman Sachs also utilized “quartiling,” in which managers assigned each employee to one of five “quartiles” until 2016 and to one of four “quartiles” from 2016 to 2018. Landman Decl. ¶ 26; Farber Report ¶ 35, ECF No. 1188-3. Managers took into account a number of criteria, including the 360 reviews and quality of performance. ECF Nos. 1258-18 at 1, 1258-19 at 3. The purpose of quartiling was to identify the top, middle, and bottom performers relative to their peers. Defs. 56.1 ¶ 22. These quartiles, along with other data points, were used to determine compensation. Pls. 56.1 ¶ 12; ECF No. 1258-18 at 1.

B. Promotion

At Goldman Sachs, promotion from Vice President to Managing Director did not involve an application process. Rather, throughout the class period, business unit heads and other managers develop lists of candidates for promotion in each of the relevant divisions. Defs. 56.1 ¶¶ 24–25, 31; Pls. 56.1 ¶ 14. Cross-ruffers, the name given to the evaluators, interview roughly a dozen people familiar with the candidates’ work, including managers, peers, and internal clients, then develop a ranked list of candidates based on the interviews. *Id.* ¶¶ 28–30. The cross-ruffers then submit this list to division heads and Goldman Sachs’ management committee. Pls. 56.1 ¶¶ 21–22.

The parties contest whether the “cross-ruffing” refers to the entire promotion process or only the second step in the process. *Compare* Pls. Summ. J. Opp’n at 26–28, ECF No. 1310, with Defs. Summ. J. Mem. at 12, 22, ECF No. 1240. In certifying the class, the Court considered

the term “cross-ruffing” to include the full promotion process—both selecting the candidates and vetting them. *See* Class Cert. Order at 8–10, 27, 30–31.

II. Procedural History

On September 16, 2010, Plaintiffs Chen-Oster, Parisi, and Orlich filed a class action alleging intentional discrimination, disparate impact discrimination, retaliation, and pregnancy discrimination claims under Title VII and the NYCHRL.³ *See* Compl. The Honorable Leonard B. Sand originally presided over this case before it was reassigned to the undersigned on May 24, 2013. ECF No. 181.

On March 30, 2018, the Court certified a class consisting of female Associates and Vice Presidents employed in the United States by Goldman Sachs and its predecessors in three of the revenue-generating divisions—Investment Banking, Investment Management, and Securities—who were subject to 360 review, quartiling, or cross-ruffing, pursuant to Federal Rule of Civil Procedure 23(b)(3) (the “Class Certification Order”). *See* Class Cert. Order at 22–49. The class includes (1) female Associates and Vice Presidents in the three divisions who were subject to either 360 review, quartiling, or both 360 review and quartiling from July 7, 2002 for those based in New York City and from September 10, 2004 for all other U.S.-based individuals through the resolution of this action, and (2) female Vice Presidents from the three divisions who were subject to the cross-ruffing process during the same time periods. *See id.* The Court concluded that certification was warranted because Plaintiffs had demonstrated that Defendants had employed a “common mode of exercising discretion,” through the three processes, to support a disparate impact class. Class Cert. Order at 24–28, 41 (citation omitted). The Court found that class resolution was also appropriate for Plaintiffs’ disparate treatment claim, which relied on the

³ Plaintiffs Gamba and De Luis were added in the second amended complaint and Plaintiff Parisi was removed. *See* Second Am. Compl.

same statistical evidence.⁴ *Id.* at 41, 45–47. On July 12, 2021, the parties’ filed motions to exclude expert testimony. ECF Nos. 1184, 1187. On July 22, 2021, Defendants filed their motion for decertification of the class. ECF No. 1194. Finally, on August 9, 2021, the parties filed cross-motions for summary judgment. ECF Nos. 1239, 1247. The Court addresses each in turn.

DISCUSSION⁵

I. Expert Testimony

The parties cross-move to exclude the opinions, reports, and testimony of certain experts under Federal Rule of Evidence 702. ECF Nos. 1184, 1187.

A. Legal Standard

The admissibility of expert testimony is governed by Federal Rule of Evidence 702. The rule provides, in relevant part, that “[a] witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion” if:

- (a) the expert’s scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
- (b) the testimony is based on sufficient facts or data;
- (c) the testimony is the product of reliable principles and methods; and
- (d) the expert has reliably applied the principles and methods to the facts of the case.

Fed. R. Evid. 702.

⁴ The Court did not certify the disparate treatment “boy’s club” claim, which would have used anecdotal evidence of a “boy’s club” culture to demonstrate disparate treatment. Class Cert. Order at 47–49.

⁵ Portions of the briefs, expert reports, and other documents discussed in this order were filed under seal or redacted. These materials are “judicial documents,” as they are “relevant to the performance of the judicial function and useful in the judicial process.” *Lugosch v. Pyramid Co. of Onondaga*, 435 F.3d 110, 119 (2d Cir. 2006); *see also Bernstein v. Bernstein Litowitz Berger & Grossmann LLP*, 814 F.3d 132, 139 (2d Cir. 2016). To the extent that information in those documents is disclosed in this order, the privacy and business interests that justified their sealing or redaction are outweighed by “the public’s right of access to [information] necessary to understand the basis for court rulings.” *Spinelli v. Nat'l Football League*, 903 F.3d 185, 193 n.2 (2d Cir. 2018); *see also Dodona I, LLC v. Goldman, Sachs & Co.*, 119 F. Supp. 3d 152, 155 (S.D.N.Y. 2015).

District courts are the gatekeepers of expert testimony, responsible for “ensuring that an expert’s testimony both rests on a reliable foundation and is relevant to the task at hand.”

Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579, 597 (1993). “[T]he proponent of expert testimony has the burden of establishing by a preponderance of the evidence that the admissibility requirements of Rule 702 are satisfied.” *United States v. Williams*, 506 F.3d 151, 160 (2d Cir. 2007).

Courts first address “the threshold question of whether a witness is ‘qualified as an expert by knowledge, skill, experience, training, or education’ to render his or her opinions.” *Nimely v. City of New York*, 414 F.3d 381, 396 n.11 (2d Cir. 2005) (quoting Fed. R. Evid. 702). After considering the expert’s qualifications, courts determine whether the expert testimony is reliable. In determining reliability, the court considers “the theory’s testability, the extent to which it ‘has been subjected to peer review and publication,’ the extent to which a technique is subject to ‘standards controlling the technique’s operation,’ the ‘known or potential rate of error,’ and the ‘degree of acceptance’ within the ‘relevant scientific community.’” *United States v. Romano*, 794 F.3d 317, 330 (2d Cir. 2015) (quoting *Daubert*, 509 U.S. at 593–94). “A minor flaw in an expert’s reasoning or a slight modification of an otherwise reliable method will not render an expert’s opinion *per se* inadmissible.” *Amorgianos v. Nat’l R.R. Passenger Corp.*, 303 F.3d 256, 267 (2d Cir. 2002). Rather, “[t]he judge should only exclude the evidence if the flaw is large enough that the expert lacks good grounds for his or her conclusions.” *Id.* (quotation marks and citation omitted). “The mere fact that an expert’s testimony conflicts with the testimony of another expert or scientific study does not control admissibility.” *In re Zyprexa Prods. Liab. Litig.*, 489 F. Supp. 2d 230, 285 (E.D.N.Y. 2007).

Even after qualifying a witness as an expert, and determining that the opinion is reliable, courts ask “whether the expert’s testimony (as to a particular matter) will assist the trier of fact.” *Nimely*, 414 F.3d at 397 (internal quotation marks and citations omitted). The testimony must be relevant, and it should not be “directed solely to lay matters which a jury is capable of understanding and deciding without the expert’s help.” *Arista Records LLC v. Lime Grp. LLC*, No. 06 Civ. 5936, 2011 WL 1674796, at *4 (S.D.N.Y. May 2, 2011) (quoting *United States v. Mulder*, 273 F.3d 91, 104 (2d Cir. 2001)).

Although the proponent of expert testimony bears the burden of establishing by a preponderance of the evidence that the admissibility requirements of Rule 702 are met, *Williams*, 506 F.3d at 160, the district court is the ultimate gatekeeper, *see Fed. R. Evid. 104(a)*. Exclusion of expert testimony is “the exception rather than the rule.” *Media Glow Digital, LLC v. Panasonic Corp. of N. Am.*, No. 16 Civ. 7907, 2019 WL 1055527, at *1 (S.D.N.Y. Mar. 6, 2019) (quoting Advisory Committee Notes to the 2000 Amendments to Fed. R. Evid. 702).

B. Plaintiffs’ Experts

Defendants move to exclude the testimony, in part or in whole, of four experts proffered by Plaintiffs. Defs. Daubert Mot., ECF No. 1187. The Court shall address each expert in turn.⁶

1. Caren Goldberg, Ph.D.

Plaintiffs offer the testimony of Caren Goldberg, Ph.D., on “Defendant Goldman Sachs’ policies, practices, and procedures with respect to diversity and inclusion and on whether its climate is tolerant of gender bias and harassment.” Goldberg Report ¶ 1, ECF No. 1188-4. Her report concludes that Goldman Sachs “maintained a climate tolerant of bias and harassment” and

⁶ Plaintiffs have represented that the opinion of David Yermack, Ph.D., need not be addressed if the opinion of Brian Dunn is excluded. *See* Pls. Daubert Opp’n at 1, ECF No. 1262. Because the Court excludes Dunn’s opinion, *see infra* Part I.C.4, the Court DENIES as moot Defendants’ motion to exclude Yermack’s opinion.

failed to meaningfully address diversity at the firm. *Id.* ¶ 13. Defendants do not contest Goldberg’s qualifications; rather, they seek to exclude her testimony on the grounds that her opinion is irrelevant to the certified claims, is methodologically flawed, and is not within the ambit of an expert opinion because it functions as a legal brief and opines on matters a jury is capable of understanding without the aid of an expert. Defs. Daubert Mem. at 10, ECF No. 1188.

For an opinion to be admissible, it must be “sufficiently tied to the facts of the case that it will aid the jury in resolving a factual dispute.” *E.E.O.C. v. Bloomberg*, No. 07 Civ. 8383, 2010 WL 3466370, at *6 (S.D.N.Y. Aug. 31, 2010) (quoting *Daubert*, 509 U.S. at 591). Expert evidence is not immune from the relevance requirement of Federal Rule of Evidence 401. *United States v. Khan*, 787 F.2d 28, 34 (2d Cir. 1986). Defendants first seek to exclude Goldberg’s testimony because it references Goldman Sachs’ policies that Defendants contend are relevant only to the uncertified “boy’s club” theory of disparate treatment. Defendants argue, therefore, that Goldberg’s opinion is irrelevant to the certified liability claims. Defs. Daubert Mem. at 11–13; Class Cert. Order at 45–47. Although the Court did not certify claims based on anecdotal evidence, the Court recognizes that at trial Defendants may seek to present evidence related to their policies, practices, and procedures with respect to diversity and inclusion to rebut Plaintiffs’ disparate treatment claims. See *United States v. City of New York (City of New York III)*, 717 F.3d 72, 84–86 (2d Cir. 2013). If Defendants present this evidence, they will put these, and other related policies, at issue. *Id.* In this scenario, Goldberg’s opinion will likely be relevant because it would contradict the evidence adduced by Defendants. See June 25, 2019 Order at 1–2, ECF No. 767; Oct. 7, 2019 Order at 7–8, ECF No. 873. Therefore, if Defendants

put their policies at issue at trial, Goldberg’s opinion shall be admissible to the extent it is relevant to respond to that evidence.

Defendants also argue that Goldberg’s opinion should be excluded because it will not aid the jury’s understanding of the facts. Defs. Daubert Mem. at 13–14, 18–19. The Court disagrees. Although experts may not reach the ultimate issue or “function[] as little more than a legal brief that parrots [the] plaintiffs’ arguments,” *Choi v. Tower Research Cap. LLC*, 2 F.4th 10, 20 (2d Cir. 2021), they may provide context to the jury to aid its understanding of the facts, *United States v. Brown*, 776 F.2d 397, 400–01 (2d Cir. 1985). In her report, Goldberg provides frameworks to aid the jury in “understand[ing] the evidence better and provid[es] [it] with the necessary tool[s] to make [the] ultimate determination.” *Fin. Guar. Ins. Co. v. Putnam Advisory Co.*, No. 12 Civ. 7372, 2020 WL 4251229, at *4 (S.D.N.Y. Feb. 19, 2020) (citation omitted). The average juror may know about diversity and inclusion policies; however, they likely lack a framework to analyze employment policies and their implementation, and to determine if the policies are effective in improving diversity and inclusion. Goldberg will assist the jury in evaluating Defendants’ institutional culture and climate, which may be relevant to the disparate treatment claims. See *City of New York III*, 717 F.3d at 83–85.

Defendants also seek to exclude Goldberg’s testimony because she relies on materials provided by Plaintiffs’ counsel at Goldberg’s request. See Defs. Daubert Mem. at 14–15; Goldberg Dep. at 142–150, ECF No. 1263-5. It is logical that experts would receive materials from counsel as they would have no other way to access non-public documents produced in discovery. Indeed, Defendants similarly provided their experts with requested materials. See, e.g., Dunn Dep. at 121, ECF No. 1263-12. Defendants present no evidence that Plaintiffs’ attorneys cherry-picked the information given to Goldberg or did not give her all the information

she requested. Defs. Daubert Mem. at 15. And, to the extent Defendants fault Goldberg for not crediting certain documents that may contradict her opinions, *id.* at 16–17, those critiques go to the weight of the opinion, not its admissibility, *Daniels v. City of New York*, No. 16 Civ. 9080, 2018 WL 5919307, at *5 (S.D.N.Y. Nov. 13, 2018).

Finally, Defendants contend that Goldberg’s methodology does not comport with best practices because she did not retain notes about how she selected information for inclusion in her report. Defs. Daubert Mem. at 17–18. Contrary to Defendants’ assertion, Goldberg kept a record of all of the materials she reviewed. Goldberg Report at Apps. 2–4; Goldberg Rebuttal at App. A, ECF No. 1188-13. She did not, however, retain the excerpts of the materials that she pulled to potentially include in her report, which the parties refer to as the “second order codes.” Goldberg Dep. at 174–75. Although she did not retain the second order codes, Goldberg’s case study conforms to the acceptable norms of her scientific community, which has not focused on the replicability of research. *See* Herman Aguinis & Angelo M. Solarino, *Transparency & Replicability in Qualitative Research*, 40 Strategic Mgmt. J. 1291, 1292 (2019), ECF No. 1263-15. Finally, by maintaining records of all the materials she reviewed, Goldberg adhered to the expectations of experts in this district. *Cf. LVL XIII Brands, Inc. v. Louis Vuitton Malletier S.A.*, 209 F. Supp. 3d 612, 644–45 (S.D.N.Y. 2016) (excluding expert for not maintaining a record of materials reviewed).

Accordingly, Defendants’ motion to preclude Goldberg from testifying is DENIED.

2. Sarah Butler

Plaintiffs offer the testimony of Sarah Butler, a Managing Director at NERA Economic Consulting, to rebut the testimony of Margaret Stockdale, Ph.D., an expert for Defendants who analyzed Goldman Sachs’ People Survey (the “People Survey”), a biannual employee experience

survey. Butler Report ¶¶ 1, 8–9, 19, ECF No. 1188-11. Specifically, Butler was asked to review Stockdale’s conclusion that the People Survey “shows that women believed Goldman Sachs has an inclusive and collaborative culture, often have similar perceptions as men, and do not perceive Goldman Sachs to have a hostile or biased climate.” *Id.* ¶ 8. Butler opines that Stockdale’s “conclusions are unreliable, not supported by the data, and rely on survey instruments that are biased and not fit for purpose.” *Id.* ¶ 10. Defendants move to exclude portions of Butler’s expert testimony on the ground that she is not qualified to opine on the culture of Goldman Sachs because she is an expert on survey design and sampling processes, and not an expert on assessing workplace culture or gender discrimination. Defs. Daubert Mem. at 20–22. They further allege that her methodology is unreliable and not replicable. *Id.* at 22–28.

The Court rejects Defendants’ argument that Butler’s conclusions are separate opinions on the culture of Goldman Sachs that are detached from her analysis of the surveys, *see id.* at 21. Clearly, Butler’s conclusions flow from her analysis of the surveys, *see In re Rezulin Prod. Liab. Litig.*, 369 F. Supp. 2d 398, 426 (S.D.N.Y. 2005), and are within her expertise of survey design and sampling processes, Butler Report ¶¶ 1–4. For example, Defendants take issue with Butler’s statement: “These data suggest that over time and across questions[,] there are patterns indicating women are less likely to believe they have been fairly treated with respect to compensation.” Defs. Daubert Mem. at 21 (quoting Butler Report ¶ 44) (emphasis omitted). This phrasing is a common way to report statistical findings⁷; it does not represent an opinion beyond Butler’s expertise. Defendants also take issue with Butler’s statement that the data in Goldman Sachs’ Exit Survey (the “Exit Survey”), a survey given to departing employees, may be biased because individuals may fear repercussions from airing grievances. *See id.* (quoting

⁷ See, e.g., University of Washington Psychology Writing Center, *Reporting Results of Common Statistical Tests in APA Format*, https://psych.uw.edu/storage/writing_center/stats.pdf (last visited Mar. 17, 2022).

Butler Report ¶ 59). Again, this opinion is within Butler’s expertise regarding survey design because it goes to possible bias created by the design of the Exit Survey. *See* Butler Report ¶ 59. As Defendants concede that Butler is qualified on the issues of survey design and statistical analysis, and her experience in this area confirms her qualifications, her statements regarding her analysis and opinions on the design and findings of the People Survey and the Exit Survey are within her area of expertise, and the Court shall not exclude them.

Defendants argue that Butler’s content analysis,⁸ which looks at the frequency at which certain terms are mentioned by female employees and seeks to show that “some male employees may have been adverse to diversity efforts,” Butler Report ¶¶ 62–67, should be excluded on the ground that her methodology is flawed, *Defs. Daubert Mem.* at 22–28. The Court agrees. Butler does not provide a sufficient explanation as to how she decided on the terms used in the content search, and, more importantly, why she classified the terms into the different categories. For example, she does not explain why “responsibility” is in the compensation category rather than under promotions, experience, or performance. *See* Butler Report, Ex. C. Butler’s failure to offer an explanation is exacerbated by the lack of a verification mechanism to ensure that the terms were properly categorized. *See* Butler Dep. at 226–28, ECF No. 1188-22. Absent such verification, the Court cannot conclude that her methodology is reliable because it is not clear whether the terms are related to the categories into which she placed them and, therefore, it is not clear if the data supports her conclusions. *See Riegel v. Medtronic, Inc.*, 451 F.3d 104, 127 (2d Cir. 2006).

⁸ “Content analysis is a research method that uses a set of procedures to make valid inferences from text.” Robert P. Weber, *Basic Content Analysis* 9 (2d ed., 1990). The analysis involves categorizing words. *Id.* at 12. For the inference to be valid, the “classification procedure [must] be reliable in the sense of being consistent” and must “measure or represent what the investigator intends it to measure.” *Id.*

Accordingly, Defendants' motion to preclude Butler from testifying as to paragraphs 62 through 67 of her report is GRANTED, and Defendants' motion to preclude Butler from testifying as to the remainder of her report is DENIED.

3. Henry Farber, Ph.D.

Plaintiffs offer the opinion of Henry Farber, Ph.D., an economics professor at Princeton University, to analyze whether the 360 review, quartiling, and cross-ruffing processes impacted women differently from their male counterparts and to explain if this impact accounts for compensation and promotion differences between men and women. Farber Report ¶¶ 1, 6. Farber examines this question by running multiple regression analyses that looked at the way these policies affected compensation by gender. *See id.* ¶¶ 53–57, 60. Defendants move to exclude all of Farber's expert testimony. Defs. Daubert Mot. They do not question his qualifications, but challenge his methodological decisions and the relevance of his opinions for Phase I of this case.⁹ Defs. Daubert Mem. at 36–37. The Court shall address each of Defendants' arguments in turn.¹⁰

a. Aggregation Across Divisions

Defendants first argue that Farber's analyses are fatally flawed because he does not disaggregate the data used in his regression analysis by division and run separate models for each division. *Id.* at 38–44. They contend that because the divisions interpret and implement aspects

⁹ As the Court has explained, this case shall proceed in two phases. “Phase I [shall] be devoted to generalizable proof of the alleged discriminatory impact and treatment of women through [the] three practices.” June 25, 2019 Order at 1–2; Oct. 7, 2019 Order at 2. Phase II shall address issues affecting individual class members. June 25, 2019 Order at 2. This type of bifurcation between liability and remedial phases is commonly used in class actions alleging discrimination. *See Reynolds v. Barrett*, 685 F.3d 193, 203 (2d Cir. 2012).

¹⁰ Defendants, discussing a different expert, mention in a footnote that Farber's analysis of internal transfers is unreliable. Defs. Daubert Mem. at 32 n.21. They do not make any argument regarding internal transfers in relation to Farber beyond this lone footnote. The Court shall not address this argument, which was not properly presented to the Court. *See Phoenix Light ST Limited v. Bank of New York Mellon*, No. 14 Civ. 10104, 2017 WL 3973951, at *22 n.41 (S.D.N.Y. Sept. 7, 2017) (“An argument mentioned only in a footnote is not adequately raised” (quoting *Shah v. Wilco Sys., Inc.*, 76 Fed. Appx. 383, 384 (2d Cir. 2003))).

of the compensation and promotion processes differently, aggregation is improper. *Id.* at 39–41. They further argue that Farber should have tested his decision not to disaggregate, and absent such testing, Farber’s opinion is “insufficiently ‘scientific’ to be admissible under Rule 702.” *Id.* at 42 (citation omitted).

Farber admits that there is sufficient data to run the models disaggregated by division for the 360 review and quartiling processes. Farber Dep. at 178–180, ECF No. 1225-59. Because there is no solid statistical reason for the decision to pool or not pool the data, the expert must make a judgment call based on methodological considerations, which the expert must justify with a “methodological explanation for why” their decision is appropriate. *See Reed Const. Data Inc. v. McGraw-Hill Cos., Inc.*, 49 F. Supp. 3d 385, 406–07 (S.D.N.Y. 2014), *aff’d*, 638 F. App’x 43 (2d Cir. 2016). There is no requirement that an expert run a specific analysis, so long as they have a methodological explanation for the analysis they chose. *Id.*; *see also Taylor v. District of Columbia Water & Sewer Auth.*, 241 F.R.D. 33, 43 (D.D.C. 2007).

Farber provides a valid, reasonable explanation for his modeling decision. In his deposition, he stated that looking at each division in isolation “was not the goal of [his] analysis.” Farber Dep. at 177–78. Because he “was analyzing a company-wide policy,” he “chose not to do it separately by division.” *Id.*; *see also id.* at 180, 378–380; Farber Rebuttal ¶¶ 45–48, 76, ECF No. 1188-12. Moreover, Plaintiffs claim that the processes are the same across the three divisions, Pls. Daubert Opp’n at 5–7, ECF No. 1262, a conclusion Farber agrees with based on his review of the data, *see* Farber Rebuttal ¶ 48. Therefore, Farber’s decision not to disaggregate, does not render the opinion “unreliable” or methodologically flawed such that exclusion is appropriate. *Romano*, 794 F.3d at 330.

b. Cross-ruffing

Defendants argue that Farber’s opinion on cross-ruffing should be excluded because he looks at the full Vice President-to-Managing Director promotion process, rather than only the evaluation step in which the cross-ruffers assess the list of candidates for promotion. Defs. Daubert Mem. at 47–49. Defendants further contend, that, even if Farber assesses both steps, he should have disaggregated his analysis to consider the nomination and interview stages separately. *Id.* The Court finds both arguments unpersuasive.

To make a claim relying on statistical evidence, a plaintiff must identify a “specific discriminatory employment practice.” *See Chin v. Port Auth. of N.Y. and N.J.*, 685 F.3d 135, 154 (2d Cir. 2012). But, Title VII expressly provides that, if the plaintiff can demonstrate to the Court “that the elements of a [defendant’s] decisionmaking process are not capable of separation for analysis, the decisionmaking process may be analyzed as one employment practice.” 42 U.S.C. § 2000e-2(k)(1)(B)(i). “Whether a particular decisionmaking process is capable of separation for analysis largely turns on the details of the specific process and its implementation in a given case.” *Chin*, 685 F.3d at 154. Here, Plaintiffs have met their burden of showing that the promotion process is a single employment practice as a matter of law.

First, the subcomponents Defendants propose are not particular employment practices but rather two interwoven steps of one singular promotion practice. *See id.* at 154–55. It is the process, not the individual steps, that determines whether someone receives a promotion. *Id.* (describing a similar promotion process). As noted above, the Court’s prior decisions indicate that it understands cross-ruffing to be a single promotion process that encompasses both nomination and evaluation by cross-ruffers, which together leads to the ultimate promotional decisions. *See Class Cert. Order at 8–10* (“[T]he vetting process for promoting Vice Presidents

to Managing Directors is known as ‘cross-ruffing.’” (citations omitted)). The Court considers this promotional process as one of the employment practices at issue in this case. *See id.*

Second, Plaintiffs have demonstrated that separate analysis would yield inadequate sample sizes. Farber Rebuttal ¶ 81. Promotion to Managing Director is rare, with an average of 73 promotions per year. Shaw Report ¶ 224, ECF No. 1192-9.¹¹ Only 204 women were promoted to Managing Director from 2003 to 2018. Farber Report ¶¶ 125, 132–133. The subcomponents are, therefore, not capable of separate analysis because of the small sample size at issue. *See* Farber Rebuttal ¶ 81. Thus, the Court finds that Farber articulates a “solid statistical reason for [the] decision to pool the data” regarding the different subcomponents of the cross-ruffing process and shall not exclude his opinion on this basis. *Reed Const. Data Inc.*, 49 F. Supp. 3d at 407.

c. 360 Review from 2002 to 2004

Defendants argue that Farber’s analysis of the 360 review process from 2002 to 2004 should be excluded because of the small sample size. Defs. Daubert Mem. at 49–50. The Court agrees. The 360 review process began producing data in 2003, Farber Report ¶ 45, so there is no data for 2002. For 2003 and 2004, there were only seventeen 360 reviews from two of the divisions. ECF No. 1188-30. Absent the use of oversampling,¹² there is insufficient data in two of the divisions for 2003 and 2004 to allow a classwide analysis, as any analysis in those two years will over-represent the one division with sufficient data. Exclusion is appropriate when an expert does not show that his data is “a fair proxy” for the process he claims to be analyzing,

¹¹ Defendants put forward Kathryn Shaw, Ph.D., as a rebuttal expert to Farber. Shaw Report ¶ 19.

¹² Oversampling is a “sampling strategy in which certain subsets of participants are overrepresented in a study group compared to the larger population in which they are drawn.” Am. Psych. Assoc., Oversampling, <https://dictionary.apa.org/oversampling> (last visited Mar. 17, 2022). Oversampling is done to ensure that certain groups, in this case data from certain years, that are small in size are adequately represented in the analysis. *See Gulino v. Bd. of Educ. of City Sch. Dist. (Gulino IV)*, 113 F. Supp. 3d 663, 679–81 (S.D.N.Y. 2015).

here, the 360 reviews of all three divisions. *See Apple v. Atl. Yards Dev. Co.*, No. 11 Civ. 5550, 2015 WL 11182422, at *6, 9–10 (E.D.N.Y. Mar. 31, 2015); *see also* Fed. R. Evid. 702(b) (an opinion must be “based on sufficient facts or data”). The Court shall, therefore, preclude Farber from offering his opinions on the 360 review process for 2002 to 2004.¹³

d. Causation and Residuals

Defendants next move to exclude Farber’s opinion that the residuals¹⁴ of his regression analyses are “sex discrimination.” Defs. Daubert Mem. at 50–51. They object to that label both on statistical grounds and as an improper legal conclusion. *Id.*

Defendants contend that Farber’s regression analyses are flawed because he excludes from his analyses production variables, which attempt to capture an individual’s productivity in a manner relevant to their role. *Id.* Farber explains that he excludes production variables because he suspected that they may be affected by gender bias, as the production variables are dependent on assignments controlled by Defendants, and, therefore, using them could introduce bias into the analyses. Farber Rebuttal ¶ 53. Additionally, because data on productivity is not available across the class or across divisions, and is often reported at the group, rather than the individual level, Farber reasons that inclusion of the variables would not truly compare employees on the

¹³ Plaintiffs argue that Defendants are precluded from making arguments about the lack of data for 2002 to 2004 because of an agreement made in 2012 permitting Defendants not to produce data from an old database. Pls. Daubert Opp’n at 14–15. Because Defendants agreed not to contest the data for 2002 to 2004 only “for the purposes of class certification,” ECF No. 159 at 40, the Court does not consider this agreement relevant here.

¹⁴ A residual in a regression analysis is “the difference between the value of an empirical observation and the value predicted by a model.” Am. Psych. Assoc., Residual, <https://dictionary.apa.org/residual> (last visited Mar. 17, 2022). In employment discrimination cases, “[b]y identifying those legitimate criteria that affect the decision[-] making process, [] plaintiffs can make predictions about what job or job benefits similarly situated employees should ideally receive, and then can measure the difference between the *predicted* treatment and the *actual* treatment of those employees. If there is a disparity between the predicted and actual outcomes for female employees, plaintiffs in a disparate treatment case can argue that the net ‘residual’ difference represents the unlawful effect of discriminat[ion] [] on the allocation of jobs or job benefits.” *Ottaviani v. State Univ. of N.Y. at New Paltz*, 875 F.2d 365, 367 (2d Cir. 1989) (emphasis in original). Specifically, “[b]y accounting for all of the ‘legitimate’ factors that could affect [compensation] and [promotions] in general, the plaintiffs hope[] to prove that there was a net ‘residual’ difference or disparity between the predicted and actual [compensation] and [promotion] of female [employees] that could only be attributed to ongoing gender discrimination [by defendants].” *Id.*

issue of productivity. *Id.* ¶¶ 55, 57. Although production may be a variable worthy of consideration, its exclusion does not render Farber's analysis unacceptable, absent some other infirmity. *Bazemore v. Friday*, 478 U.S. 385, 400–01 (1986) (Brennan, J., joined by all other Members of the Court, concurring in part). As Farber's decision to exclude production variables is well-reasoned and there are no other fatal infirmities, his regression analyses are admissible. *Id.* His choice to exclude the production variables goes to the weight of his testimony rather than its admissibility. *Reed Const. Data Inc.*, 49 F. Supp. 3d at 401.

Moreover, Farber's finding that the residuals represent gender-based differences is also founded. The use of statistics in sex discrimination cases is common. *See United States v. City of New York (City of New York II)*, 731 F. Supp. 2d 291, 300 (E.D.N.Y. 2010) (disparate impact); *City of New York III*, 717 F.3d at 85 (disparate treatment). Indeed, the Second Circuit has stated that “[i]t is apodictic that ‘statistical reports may . . . be admissible to support an inference of discrimination.’” *Bickerstaff v. Vassar College*, 196 F.3d 435, 448 (2d Cir. 1999) (quoting *Raskin v. Wyatt Co.*, 125 F.3d 55, 67 (2d Cir. 1997)). Regression modeling, the form of statistical analysis Farber used, attempts to eliminate other possible causes, leaving only the studied variable as the cause of the difference observed, also called the residual. *Ottaviani*, 875 F.2d 365 at 367. In employment discrimination cases, statistical significance at the 5% level, or approximately two standard deviations, is sufficient to make a *prima facie* case. *Smith v. Xerox Corp.*, 196 F.3d 358, 366 (2d Cir. 1999), *overruled on other grounds by Meacham v. Knolls Atomic Power Lab.*, 461 F.3d 134 (2d Cir. 2006). Such a statistical significance means that there is a 5% probability that the difference found is due to chance rather than being caused by the studied process. *Id.* Here, Farber's regression analyses show statistical significance at the 5% level, so his studies indicates that there is a 5% chance that the residuals (*i.e.* the difference in

treatment) result from something other than differences in the employee’s gender. Farber Report ¶¶ 76, 79, 81–83, 98–102, 104–105, 134–135. The Court, therefore, finds that Farber can testify that the differences he found were based on gender.

Defendants next argue that, because Farber was not prepared to make a conclusion about gender discrimination at the class certification stage, his “about face” should be questioned. Defs. Daubert Mem. at 50–51. The Court disagrees. Farber’s opinion is methodologically sound and, therefore, admissible under Federal Rule of Evidence 702. The fact that Farber was not prepared to state that the residuals represented gender-based differences at the class certification hearing does not bar him from making that conclusion now, after he has had the opportunity to update his model to include information ascertained through discovery.

Furthermore, Defendants argue that Farber’s opinion represents an impermissible legal conclusion. *Id.* at 51. “An opinion is not objectionable just because it embraces an ultimate issue.” Fed. R. Evid. 704(a). It is permissible for Farber to opine that the residuals of his models represent the gender-based difference because he sufficiently controlled for other plausible causes of the difference. Even though the opinion embraces the ultimate issue, it does not preclude the jury from rejecting the opinion. However, to the extent that Farber uses the term “sex discrimination,” which is a legal term of art, he is making a legal conclusion that is improper as it usurps the role of the jury. *In re Platinum-Beechwood Litig.*, 469 F. Supp. 3d 105, 115 (S.D.N.Y. 2020). Farber may opine on the residuals and their possible meaning so long as he does not use the term “sex discrimination.”

e. Damages and Compensation

Defendants move to exclude Farber’s analysis of damages and compensation that is not tied to the analysis of the three processes at issue. Defs. Daubert Mem. at 45–47, 52–55. The

Court has stated that “compensation may be a topic addressed in Phase I insofar as it is connected to one or more of the three challenged processes.” June 25, 2019 Order at 1–2; *see also* Oct. 7, 2019 Order at 5–6. But, damages are beyond the scope of the Phase I trial, which shall focus on generalizable issues of liability. *See* ECF No. 657 at 2 (stating damages are part of Phase II); ECF No. 888 at 3.

For an opinion to be admissible, it must be “sufficiently tied to the facts of the case that it will aid the jury in resolving a factual dispute.” *Bloomberg*, 2010 WL 3466370, at *6 (citation omitted). To the extent that Farber’s opinion looks at compensation generally to ascertain how the processes at issue impacted compensation, compensation is relevant, and the Court shall not exclude those portions of Farber’s opinion. *See* Oct. 7, 2019 Order at 5–7. To the extent that Farber’s opinion addresses a historical pay theory whereby compensation is reduced because of past use of the processes, it goes to damages, rather than liability, and thus, is properly excluded as irrelevant to the Phase I trial. *See Bloomberg*, 2010 WL 3466370, at *6. Moreover, to the extent that Farber’s models are not related to the challenged processes but rather look to show general pay disparities for damages purposes, they are not relevant to Phase I and are, therefore, excluded. *See* Farber Report ¶¶ 136–57; Farber Rebuttal Report ¶¶ 108–19.

Accordingly, Defendants’ motion to preclude Farber from testifying as to the 360 review process for 2002 to 2004 and damages is GRANTED, and Defendants’ motion to preclude Farber from testifying as to the remainder of his reports is DENIED.

C. Defendants’ Experts

Plaintiffs move to exclude the opinions of four experts put forward by Defendants. Pls. Daubert Mot., ECF No. 1184. The Court shall address each expert in turn.

1. Kathryn Shaw, Ph.D.

Defendants offer Kathryn Shaw, Ph.D., an economics professor at the Stanford Graduate School of Business, to rebut Farber’s opinion that the challenged processes resulted in gendered differences in compensation and promotions. Shaw Report ¶¶ 1, 19. Shaw has an extensive background in economics research and modeling, focusing on labor economics. *Id.* ¶¶ 2–14. Defendants asked Shaw to review Farber’s report, provide a critique, and create her own models to test Plaintiffs’ gender discrimination claims. *Id.* ¶ 21. Plaintiffs move to exclude portions of Shaw’s expert opinion. Pls. Daubert Mot. Specifically, they move to exclude Exhibits 2–9, 20–29, and 31–69, and all related text. *Id.* Plaintiffs do not challenge Shaw’s experience or knowledge; however, they do note that she lacks experience in making models for gender discrimination claims. Pls. Daubert Mem. at 5, 5 n.3, ECF No. 1191.

a. Production Variables

Plaintiffs argue that Shaw’s inclusion of production variables is unsound and the resulting exhibits and text should be excluded. *Id.* at 6. They contend that she was not involved in choosing the variables, cannot explain why certain data types were used to represent productivity, and did not ensure the variables were measuring what they purported to measure or that there was sufficient data to use the variables. *Id.* at 7–11.

For a regression analysis to be admissible under *Daubert*, there must be a “predictable and justifiable methodology,” reflecting modeling choices that are based on “learned scholarship” and are reasonably supported. *Reed Const. Data Inc.*, 49 F. Supp. 3d at 400. Expert evidence is properly excluded when there is “too great an analytical gap between the data and the opinion proffered” to satisfy Rule 702. *Laumann v. Nat’l Hockey League*, 117 F. Supp. 3d. 299, 319 (S.D.N.Y. 2015) (citation omitted). Typically, including an

“irrelevant . . . variable[] [in a regression model] will go to the probative value of the analysis, not its admissibility.” *Sobel v. Yeshiva Univ.*, 839 F.2d 18, 36 (2d Cir. 1988).

Contrary to Plaintiffs’ contentions, Shaw did choose the variables she included in her models and explains why she chose each variable. Shaw Report ¶ 74, App. C; *see, e.g.*, Shaw Dep. at 105–06, ECF No. 1297-6. Because Shaw explains the reasons for including a production variable and identifies variables for different subgroups in the three divisions, Shaw Report ¶¶ 74, 151–162, the Court considers her choices methodologically sound. The fact that there is overlap with a prior expert for Defendants does not invalidate her decision, *see Am. Home Assur. Co. v. Merck & Co.*, 462 F. Supp. 2d 435, 448 (S.D.N.Y. 2006), nor does the fact that she had research assistants aiding her in reaching the decision, *see Phoenix Light SF Limited v. Bank of New York Mellon*, No. 14 Civ. 10104, 2020 WL 1322856, at *10 (S.D.N.Y. Mar. 20, 2020).

Plaintiffs are within their right to be concerned that production variables do not exist for all individuals. *See* Farber Rebuttal ¶ 55. They further point out that Shaw’s model assumes that the production variables for those who have them are representative of those who do not. *Id.* ¶ 58. Farber tested whether there is a correlation between those missing production variables and other variables of interest. *Id.* He concludes that missing production data is correlated with gender and compensation. *Id.* This indicates that the production variables may introduce bias to the models. *Id.* ¶¶ 53, 58. Although Plaintiffs’ concerns are valid, the Court does not find that they require exclusion of Shaw’s opinion because Shaw cogently articulates her basis for including the production variables. Shaw Report ¶¶ 151–58, 161, 285. She acknowledges that the production variables differ, on average, by gender. *Id.* ¶ 156. And, although Farber views gender differences as a reason to omit the variables, Farber Rebuttal ¶ 53, Shaw views this correlation as an indication that part of the gender difference in the residual may be from not

correcting for gender differences in production rather than from the processes at issue in this case. Shaw Report ¶¶ 156–57.

Moreover, there was production data for the majority of individuals, Farber Rebuttal, Tbl. 7, and Shaw accounts for the missing data by including a missing indicator variable for each production variable, Shaw Report at App. C. The use of a missing indicator variable is an accepted way to correct for missing data under these circumstances, even though it may result in a biased analysis. *See Michael P. Jones, Indicator and Stratification Methods for Missing Explanatory Variables in Multiple Linear Regression*, 91 J. Am. Stat. Assoc. 222, 222, 228 (2019). And, so long as the decision to include or exclude a variable is reasonable, the fact that the choice may “render the analysis less probative than it otherwise might be” goes to weight rather than admissibility. *Cf. Bazemore*, 478 U.S. at 400; *see also Sobel*, 839 F.2d at 36. The missing data does not, therefore, provide a basis for the Court to exclude Shaw’s analysis.

b. Disaggregation Across Divisions

Plaintiffs next allege that Shaw’s analysis is fatally flawed and irrelevant because it disaggregates compensation and promotion data by division and business unit instead of considering the data on a classwide basis. Pls. Daubert Mem. at 16.

Disaggregation can reduce the statistical power of a model by reducing the sample size because “[statistical] power increases with sample size. A gender gap of a specific magnitude (say 15%) is more likely to show up as statistically significant in a regression analysis based on 1,000 observations than in one based on 100 observations.” William T. Bielby and Pamela Coukos, “*Statistical Dueling*” with Unconventional Weapons: What Courts Should Know About Experts in Employment Discrimination Class Actions, 56 Emory L.J. 1563, 1597 (2007).

Initially, an expert must show that any decision regarding disaggregation is methodologically sound. Fed. R. Civ. P. 702(c). Experts should avoid disaggregation when it would reduce the sample size to the point where statistical modeling is ineffective. *See* March 10, 2015 Order at 18–21, ECF No. 363. But, even when there is no statistical reason compelling disaggregation, the decision to disaggregate is not a basis for exclusion so long as the expert offers a methodological explanation for their decision. *Cf. Reed Const. Data Inc.*, 49 F. Supp. 3d at 400, 406–07.

Disaggregating to the business unit level results in sample sizes too small to produce statistically significant results. *See* Shaw Report at App. G., Exs. 59, 64. The amount of data at the business unit level ranges as high as 1,482 person years to as low as 2 person years.¹⁵ *Id.* at Apps. E, F, G. When the amount of data is small, statistical significance may be impossible to detect even if there is a causal relationship. *Chin*, 685 F.3d at 153; *see also* Shaw Dep. at 314–16. Some business units may have sufficient data to allow analysis; however, others do not. Disaggregation at the business unit level could mask differences because of the small sample sizes, *see Moussouris v. Microsoft*, 311 F. Supp. 3d 1223, 1236 (W.D. Wash. 2018); *see also* March 10, 2015 Order at 18–21, making disaggregation at the business unit level methodologically flawed. The Court, therefore, finds that Shaw’s modeling at the business level unit is methodologically unsound and should be excluded under Federal Rule of Evidence 702.

But, there is likely sufficient data to disaggregate at the division level. *See* Farber Dep. at 178–80 (testifying as to the 360 review and quartiling processes). Plaintiffs argue that because the issue in the case is whether the companywide policies discriminate against women, the

¹⁵ A “person year” is “the sum of the number of years that each individual in a population has been affected by an event, occurrence, or condition.” Am. Psych. Assoc., Person-years, <https://dictionary.apa.org/person-year> (last visited Mar. 17, 2022).

proper level of analysis is companywide, controlling for business unit and division. Pls. Daubert Mem. at 19. Shaw states that she decided to disaggregate at the division level because she believes the divisions function differently and are decentralized. Shaw Report ¶¶ 164–68. Shaw ran Chow tests¹⁶ to ascertain if unpooling the divisions was statistically sound and concluded the divisions should not be analyzed together because the Chow tests showed that the gender differences were not the same across the divisions. *Id.* ¶¶ 169–71, 179. Shaw, therefore, offers a methodologically sound reason to disaggregate the data. Fed. R. Civ. P. 702(c).

Finally, the Court considers whether the disaggregated models would be helpful to the jury and relevant to the task at hand. *Daubert*, 509 U.S. at 596–97. Because this question depends on how the jury decides to assess the differences and similarities among the various relevant divisions, the jury—not the Court—must decide if pooled or unpooled data makes more sense in light of the question asked, the data presented, and the explanations of the experts. *In re Zyprexa*, 489 F. Supp. 2d at 285. It is for the jury to decide which theory of the case they find persuasive. *Id.* (quoting *Daubert*, 509 U.S. at 596).

c. Outliers

Shaw excluded predominantly male “outliers” who earned considerably more than others out of concern that their inclusion would improperly skew her models. Shaw Report ¶¶ 253–59. Plaintiffs contend that removing high-earning outliers is methodologically flawed, especially in a gender discrimination case looking at compensation differences between men and women. Pls. Daubert Mem. at 22–24.

Shaw ran her model excluding three different groups of alleged outliers: (1) those with compensation greater than 1.5 times the interquartile range above 75%, (2) the top 5% of earners,

¹⁶ A Chow Test is a statistical test to determine if a model’s coefficients are identical, or equal, across subsets of data. Shaw Report ¶ 169 n.216; Farber Rebuttal ¶ 44.

and (3) the top 10% of earners. Shaw Report ¶ 258, Ex. 34. She concludes that “[t]he estimated gender gap become[s] more favorable to women once . . . extraordinarily high earners” are removed, in part, because most of the high earners are men. *Id.* ¶ 259.

Even if, as Defendants argue, excluding outliers may be a common practice in econometrics, Defs. Daubert Opp’n at 24, Defendants do not cite any cases condoning the practice when the models intend to look at compensation disparities and those excluded are almost entirely from the unprotected class.¹⁷ Outliers should only be removed after “careful examination” to ensure the outliers are not meaningful for the analysis. *See* Am. Bar Assoc., *Econometrics: Legal, Prac. and Tech. Issues* § 4.E.1 (2d ed. 2014). Shaw’s own analysis shows that the outliers she omitted were almost entirely men, which is certainly meaningful when the regressions at issue are designed to detect gender differences in compensation. Defendants have not met their burden of demonstrating that removing the outliers is methodologically sound in this context. Accordingly, Exhibit 34, and the related text, are excluded.

d. Cross-Ruffing

When assessing cross-ruffing, Shaw analyzes the different subcomponents of the cross-ruffing promotional process separately. Shaw Report ¶¶ 233–37. Plaintiffs argue that this analysis improperly disaggregates the subcomponents because cross-ruffing is one employment process. Pls. Daubert Mem. at 24–25. This Court has held, as a matter of law, that cross-ruffing is one employment process which involves multiple subcomponents incapable of being analyzed

¹⁷ The sources Shaw cites in her report do not address the removal of outliers in models designed to determine if there is a compensation disparity. For example, in her report she cites Joshua D. Angrist and Alan B. Krueger, *Empirical Strategies in Labor Economics*, in Handbook of Labor Economics Vol. 3, 1348–1349 (Orley Ashenfelter and David Card eds., 1999) for the proposition that she may remove the top 5% and 10% of earners. Shaw Report ¶ 258 n.290. However, in advocating for “trimming” techniques, the cited authors seek to remove “misreported” data in study samples, Angrist & Krueger at 1347. Shaw is not attempting to remove improperly reported data, and, therefore, the citation does not support her methodology.

separately. *See supra* Part I.B.3.b. Accordingly, Shaw’s subcomponent analysis of the cross-ruffing process is excluded as irrelevant and unhelpful to the jury.

Accordingly, Plaintiffs’ motion to preclude Shaw from testifying as to disaggregation at the business unit level, the impact of removing outliers, and the subcomponents of the cross-ruffing process is GRANTED, and Plaintiffs’ motion to preclude Shaw from testifying as to the remainder of her report is DENIED.

2. Margaret Stockdale, Ph.D.

Defendants offer Margaret Stockdale, Ph.D.’s, opinion on “whether Goldman Sachs has implemented scientifically-supported and other leading practices that work to provide a culture and climate that [are] not tolerant of gender bias and that [are] supportive of women and their career success.” Stockdale Report ¶¶ 6–7, ECF No. 1192-10. Defendants also ask Stockdale to evaluate Goldberg’s report. *Id.* ¶ 8. Plaintiffs move to exclude all of Stockdale’s expert testimony. Pls. Daubert Mot. Plaintiffs do not challenge her credentials to opine on Goldman Sachs’ corporate climate, they challenge her methodology. Pls. Daubert Mem. at 25–26. Plaintiffs also challenge Stockdale’s qualifications to assess survey design and survey data as well as her methodology as it pertains to the People Survey. *Id.* at 25. Stockdale is a professor of psychology at Indiana University-Purdue University, specializing in gender issues in the workplace. Stockdale Report ¶¶ 1–2. She has authored peer-reviewed articles and served as a consultant to companies on sexual harassment policies and complaint resolution. *Id.* ¶ 2.

a. Corporate Culture

Plaintiffs contend that, because Stockdale does not explicitly use a specific analytical framework to guide her analysis, her report is devoid of scientific methodology and “based entirely on her own feelings about the evidence.” Pls. Daubert Mem. at 26–27. Although

Stockdale does not identify a specific analytical framework, she testified that her work is implicitly grounded in the frameworks used by Goldberg because she is rebutting Goldberg's report. Stockdale Dep. at 96–97, ECF No. 1259-7. Additionally, Stockdale's report is grounded in the social science methodology of triangulation, Stockdale Report ¶¶ 37–39, and this methodology has long been accepted in her field, *see, e.g.*, Todd D. Jick, *Mixing Qualitative & Quantitative Methods: Triangulation in Action*, 602–611 (1979). Courts have also approved of methodologies akin to triangulation. *See, e.g.*, *Linde v. Arab Bank, PLC*, 922 F. Supp. 2d 316, 322–23 (E.D.N.Y. 2013). The Court finds that Stockdale's opinion on corporate culture is sufficiently methodologically grounded to satisfy Rule 702.

b. Surveys

Plaintiffs also argue that Stockdale's use of the People Survey data is deficient. Pls. Daubert Mem. at 29–35. First, they contend that the data from the People Survey is unreliable because, as an employer survey, the responses will be “artificial[ly] positiv[e] to avoid negative repercussions or gain favoritism.” *Id.* at 30. They also argue that the data from the Exit Survey is problematic because of a similar positive bias. *Id.* at 31–32. But, possible bias in the data underlying an expert's opinion “is a subject for cross-examination, and goes to the weight, not the admissibility, of the expert testimony.” *Auto. Ins. Co. of Hartford v. Electrolux Home Prod., Inc.*, No. 10 Civ. 0011, 2012 WL 6629238, at *2 (S.D.N.Y. Dec. 20, 2012). Plaintiffs further contend that Stockdale lacks the expertise to offer an opinion on the survey design and survey data. Pls. Daubert Mem. at 26, 30. However, Stockdale is well versed in study design and analysis as evidenced by her teaching of social science research methodology. Stockdale Report at App. A.

Plaintiffs next contend that Stockdale improperly created her own scale to evaluate the People Survey rather than using the cut-offs used internally by Goldman Sachs. Pls. Daubert Mem. at 31. Although Stockdale concedes that she created her own scale by collapsing some categories of responses, Stockdale Report ¶ 61 n.35, such collapsing is an accepted practice in her field because it allows for ease of interpretation and eliminates the effects of outliers and bimodal distributions, *see*, Melinda J. Moya & Alison L. O’Malley, *Mechanics of Survey Data Analysis*, in Employee Surveys & Sensing at 239, 246 (William H. Macey & Alexis A. Fink eds., 2020), ECF No. 1259-39. The Court does not find that this choice invalidates Stockdale’s opinion.

Plaintiffs also attack Stockdale’s conclusions because they differ from those of their expert. Pls. Daubert Mem. at 31–33. Experts reaching different conclusions is not a sufficient basis for exclusion. *See In re Zyprexa Products Liab. Litig.*, 489 F. Supp. 2d at 285.

Additionally, Plaintiffs argue that Stockdale’s opinion should be excluded because she relied on a research assistant and an outside firm, Cornerstone Research, to perform some of the analysis in her report. Pls. Daubert Mem. at 33. This argument also fails. Experts are permitted to “present the findings and conclusions of those whose work [they] supervised and that [they] could personally replicate if necessary.” *In re M/V MSC Flaminia*, No. 12 Civ. 8892, 2017 WL 3208598, at *22 (S.D.N.Y. July 28, 2017). Stockdale supervised her research assistant and Cornerstone Research and checked their work. Stockdale Report ¶ 10; Stockdale Dep. at 99, 107–11, 124. Stockdale teaches applied social science research methods, evidencing that she could replicate the work completed by Cornerstone and her research assistant, and Plaintiffs do not allege otherwise. *See* Stockdale Report ¶ 2. Moreover, contrary to Plaintiffs’ representations, Pls. Daubert Mem. at 33–34, Stockdale’s research assistant, her husband,

Michael Heck, Ph.D., is qualified to assist her. Stockdale deployed him to serve as a second coder.¹⁸ Stockdale Dep. at 123–24. Coders do not require any specific expertise—rather, they must be trained to apply the coding framework of the particular analysis.¹⁹ Heck holds a doctorate in applied experimental psychology, for which he likely received at least some training in qualitative research methods.²⁰ And, even if he did not receive specialized training, the Court is not convinced that Heck is incapable of being trained to be a coder.²¹ The Court shall not, therefore, exclude Stockdale’s report and testimony.

Accordingly, Plaintiffs’ motion to preclude Stockdale from testifying is DENIED.

3. Eric Dunleavy, Ph.D.

Defendants retained Eric Dunleavy, Ph.D., an industrial and organizational psychologist and the director of Employment and Litigation Services at DCI Consulting Group, to analyze the work activities of the Associates and Vice Presidents in the three divisions at issue in this case and the job relatedness of the three processes at issue. Dunleavy Work Analysis Report at 1, 10, ECF No. 1192-2; Dunleavy Job Relatedness Report at 1, ECF No. 1192-26. He provided two expert reports as well as a rebuttal report and a sur-rebuttal report. ECF Nos. 1192-2, 1192-3, 1192-26, 1192-27. In his work analysis report, Dunleavy was asked to opine on whether the Associates and Vice Presidents perform the same roles in each division. Dunleavy Work

¹⁸ Using a second coder is common in social science research as it allows for validation of coding, also known as intercoder reliability. See Kimberly A. Neuendorf, *The Content Analysis Guidebook*, 41, 159, 166 (2016), ECF No. 1259-29. Butler, Plaintiffs’ expert, also utilized a dual-coder methodology. Butler Dep. at 223, 241.

¹⁹ See Cliodhna O’Connor & Helene Joffe, *Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines*, 19 Int’l J. Qualitative Methods 1, 6 (2020); Johnny Saldana, *The Coding Manual for Qualitative Researchers* 28–30 (1st ed., 2009) (listing personal attributes needed for coding).

²⁰ See Stockdale and Associates, *Michael Heck, Ph.D.*, <https://www.stockdaleandassociates.com/michael-heck-ph-d/> (last visit Mar. 17, 2022). Currently, obtaining a Ph.D. in psychology from Kansas State University requires two courses in research design and analysis. See Kansas State University, *Department of Psychological Sciences Graduate Program Overview*, <https://www.k-state.edu/psych/graduate/> (last visited Mar. 17, 2022).

²¹ As little as two hours of training is sufficient to learn how to code. See Ellen Childs & Lindsay B. Demers, *Qualitative Coding Boot Camp: An Intensive Training and Overview for Clinicians, Educators, and Administrators*, J. of Teaching and Learning Res. (2018).

Analysis Report at 1. In his job relatedness report, he was asked to evaluate the job relatedness of the three processes at issue to determine if the processes are reliable and valid. Dunleavy Job Relatedness Report at 1–2. Plaintiffs move to exclude Dunleavy’s work analysis and rebuttal work analysis reports, arguing they are methodologically flawed and will not assist a jury as the reports address topics not at issue in this case. Pls. Daubert Mem. at 36–38. Plaintiffs move to preclude Dunleavy’s testimony on statistical studies five and seven in his job relatedness reports because they include the same production variables discussed above. *Id.* at 36 n.30.

a. Work Analysis

First, Plaintiffs rely on three cases to argue that Dunleavy’s methodology is flawed and his testimony must, therefore, be excluded. *Id.* at 36, 40, 43–44. But, these cases are inapposite because they relate to work analyses undertaken to create or validate a hiring tool. *See Guardians Ass’n of N.Y.C. Police Dep’t, Inc. v. Civ. Serv. Comm’n of City of N.Y.*, 630 F.2d 79, 95–96 (2d Cir. 1980); *Gulino IV*, 113 F. Supp. 3d at 676–83; *United States v. City of New York (City of New York I)*, 637 F. Supp. 2d 77, 100–105, 110–15 (E.D.N.Y. 2009). Here, the work analysis is being used to ascertain if Associates and Vice Presidents do the same work, rather than to create or evaluate an assessment for determining who should be hired. *See* Dunleavy Work Analysis Report at 1.

As the literature Plaintiffs cite emphasizes, a work analysis is not “one size fits all,” and the methods must align with the purpose of the analysis. *See* Paul R. Sackett et al., *Job and Work Analysis*, in 12 *Handbook of Psychology: Industrial and Organizational Psychology* 61, 63 (Neal W. Schmitt & Scott Highouse eds., 2d ed., John Wiley & Sons, Inc. 2013), ECF No. 1192-23. Dunleavy has shown that his work analysis is fit for the purpose for which it is offered—showing that the Associates and Vice Presidents perform different tasks within and

across divisions. Dunleavy Work Analysis Report at 1, 207, 209. He articulated his hypotheses: (1) that there is no difference in the tasks performed by Associates, and (2) that there is no difference in the tasks performed by Vice Presidents, *id.* at 17–18, and then designed a methodology to test those hypotheses, *id.* at 18–31.

The methodology Dunleavy uses is similar to one found to be reliable for a comparable purpose by another court in this district. *See Kassman v. KPMG LLP*, 416 F. Supp. 3d 252, 262–63, 269–71 (S.D.N.Y. 2018). Defendants have, therefore, met their burden of showing that Dunleavy’s methodology is sufficiently reliable to be admissible. *See id.* at 270. Plaintiffs’ concerns related to the underlying data, specifically the sample size and sampling methodology, go to weight rather than admissibility. *Reed Const. Data Inc.*, 49 F. Supp. 3d at 401. Similarly, issues related to the extrapolation of the work analysis conducted in 2019 to the class period also go to weight. *See Restivio v. Hessemann*, 846 F.3d 547, 577 (2d Cir. 2017) (“[G]aps or inconsistencies in the reasoning leading to [the expert’s] opinion . . . go to the weight of the evidence, not to its admissibility.” (quoting *Campbell ex rel. Campbell v. Metro. Prop. & Cas. Ins. Co.*, 239 F.3d 179, 186 (2d Cir. 2001))).

b. Relevance

Plaintiffs further argue that Dunleavy’s job analysis opinion is not relevant because it does not address gender differences. Pls. Daubert Mem. at 47–48. The Court agrees that Dunleavy’s analysis does not go to the heart of this case, which asks if the challenged processes impact women differently from men. However, Dunleavy’s job analysis opinion is relevant to the issue of commonality, which is before the Court in Defendants’ motion for class decertification. *See* Defs. Decert. Mem. at 28–31, ECF No. 1224. An opinion need not go to the ultimate issue to be relevant; it must only assist the trier of fact. *See In re Pfizer Inc. Sec. Litig.*,

819 F.3d 642, 661 (2d Cir. 2016). Therefore, Dunleavy’s job analysis opinion is admissible for the purpose of the decertification motion only.

c. Production Variables

Plaintiffs move to strike two statistical studies in Dunleavy’s job relatedness report and sur-rebuttal report. Pls. Daubert Mot.; Pls. Daubert Mem. at 36 n.30. Dunleavy posits that the 360 review and quartiling processes were intended to capture performance and, therefore, Dunleavy requested and used the production variables to test if the processes correlate with the production variables as part of validating the processes. Dunleavy Job Relatedness Report at 61. This is a reasoned justification for using the production variables as part of his analysis. The Court shall not exclude the statistical studies incorporating the production variables.

Accordingly, Plaintiffs’ motion to preclude Dunleavy from testifying as to the two statistical studies in his job relatedness reports and to prevent the Court from considering Dunleavy’s job analysis report when assessing Defendants’ decertification motion is DENIED.

4. Brian Dunn

Defendants asked Brian Dunn to opine on (1) “[w]hether and to what extent the jobs performed by professionals with the corporate title Associate and [Vice President] at financial services firms such as Goldman Sachs differ”²² and (2) “[w]hether and to what extent financial services firms such as Goldman Sachs need to tie compensation to performance.” Dunn Report ¶ 9, ECF No. 1192-4. Plaintiffs move to exclude Dunn’s reports, arguing that Dunn’s job variety opinion is unreliable and his opinion related to paying for performance is irrelevant. Pls. Daubert Mem. at 49–53. Dunn has worked for over 35 years as a compensation consultant, using data provided by financial services firms to benchmark their compensation as compared to

²² Like Dunleavy’s opinion, Dunn’s opinion on the jobs performed by the class members is only relevant to Defendants’ decertification motion.

their industry. Dunn Report ¶¶ 1–2. In the past, he has served as a compensation consultant for Goldman Sachs. *Id.* ¶ 4.

Dunn’s opinion is based on his experience in compensation benchmarking in the financial services industry. *Id.* ¶¶ 10, 26. According to Dunn, benchmarking firms ask their financial services clients to complete surveys, created and provided by the benchmarking firms, which ask the clients to indicate the compensation of various employees sorted by job category. *See id.* ¶¶ 18–19, 22. For information about the work performed by these various employees, Dunn relies almost exclusively on declarations by Goldman Sachs’ employees and David Stowell’s *Investment Banks, Hedge Funds, and Private Equity*, Third Edition (London: Academic Press, 2018). *See id.* ¶¶ 27–33. He connects the job categories used in benchmarking to the work performed by Goldman Sachs’ employees in a conclusory fashion, stating that classification “is necessary *because* the professionals perform different jobs; if they did not, they could all be lumped together.” *Id.* ¶ 35 (emphasis in original). Such a leap creates “too great an analytical gap between the data and the opinion proffered.” *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 146 (1997).

Moreover, Dunn’s opinion as it pertains to paying for performance, which would support Defendants’ business necessity rebuttal to Plaintiffs’ claims, would not aid the jury and is, therefore, excluded. “Testimony is properly characterized as ‘expert’ only if it concerns matters that the average juror is not capable of understanding on his or her own.” *United States v. Mejia*, 545 F.3d 179, 194 (2d Cir. 2008). Dunn’s report concludes that Goldman Sachs must pay its employees based on “the quality and value of their work” to be competitive in the financial services employment market. Dunn Report ¶¶ 14, 64. It is within a lay person’s understanding that companies pay individuals to do their jobs and need a means of assessing if those individuals

are doing their jobs—to measure their performance and productivity. It is also within a lay person's knowledge that companies need to pay competitively to attract and retain employees. An expert is not needed to opine that, in the financial services industry, these general propositions apply. Dunn's opinion will not add to the jury's understanding as his report does not discuss the specific mechanism of measuring performance commonly used in the industry or provide any other opinion for which expertise is necessary. *See id.*

Accordingly, Plaintiffs' motion to preclude Dunn's testimony is GRANTED.

II. Class Decertification

Defendants move to decertify the class on the grounds that, since 2018, there have been material factual and legal changes that undermine the certification. Defs. Decert. Mem. at 1–7. Defendants argue that the class lacks Article III standing and that the class no longer meets the Federal Rule of Civil Procedure 23 requirements of commonality, predominance, typicality, superiority, and adequacy. *See generally id.* Plaintiffs contend that the class has standing and that Defendants are recycling their prior arguments as to the Rule 23 factors despite no change in the relevant legal frameworks or facts. *See Pls. Decert. Opp'n at 1–2, ECF No. 1269.* The Court agrees.

A. Legal Standard

Once a class is certified, “Rule 23 provides district courts with broad authority at various stages in the litigation to revisit class certification determinations and to redefine or decertify classes as appropriate.” *Jacob v. Duane Reade, Inc.*, 293 F.R.D. 578, 594 (S.D.N.Y. 2013) (quoting *Wang v. Chinese Daily News, Inc.*, 709 F.3d 829, 836 (9th Cir. 2013), superseded on other grounds by 737 F.3d 538 (9th Cir. 2013)); *see Fed. R. Civ. P. 23(c)(1)(C)*. “A court may decertify a class if it appears that the requirements of Rule 23 are not in fact met[,] . . . [it] may

not disturb its prior findings absent some significant intervening event or a showing of compelling reasons to reexamine the question.” *Mazzei v. Money Store (Mazzei I)*, 308 F.R.D. 92, 106 (S.D.N.Y. 2015), *aff’d, Mazzei v. Money Store (Mazzei II)*, 829 F.3d 260 (2d Cir. 2016) (quotation marks and citations omitted). Such compelling reasons may include “an intervening change of controlling law, the availability of new evidence, or the need to correct a clear error or prevent manifest injustice.” *Id.* (quoting *Gulino v. Bd. of Educ. (Gulino II)*, 907 F. Supp. 2d 492, 504 (S.D.N.Y. 2012), *aff’d, Gulino v. Bd. of Educ. (Gulino III)*, 555 Fed. Appx. 37 (2d Cir. 2014)); *see also Stinson v. City of New York*, No. 10 Civ. 4228, 2014 WL 4742231, at *1 (S.D.N.Y. Sept. 23, 2014) (collecting cases). Absent such a showing, “the factual underpinnings of a court’s prior certification order are deemed to be law of the case.” *Stinson*, 2014 WL 4742231, at *2 (quotation marks and citation omitted).

“Decertification is an ‘extreme step.’” *Gulino II*, 907 F. Supp. 2d at 504 (quoting *Woe v. Cuomo*, 729 F.2d 96, 107 (2d Cir. 1984)). “A defendant seeking to decertify a class ‘bear[s] a heavy burden to prove the necessity of either the drastic step of decertification or the less draconian but still serious step of limiting the scope of the class.’” *Id.* (quoting *Gordon v. Hunt*, 117 F.R.D. 58, 61 (S.D.N.Y. 1987)). However, Plaintiffs retain “the burden [of] demonstrat[ing] that the [Rule 23(b)(3)] requirements [are] satisfied.” *Mazzei II*, 829 F.3d at 270. These requirements are numerosity, commonality, typicality, adequacy of representation, predominance of common questions of law or fact, and the superiority of a class action to other procedural mechanisms. Fed. R. Civ. P. 23(a), (b)(3).

B. Analysis

1. Standing

Defendants argue that each class member must demonstrate standing to maintain certification for a Rule 23(b)(3) damages class action. Defs. Decert. Mem. at 1–3, 23–27.²³ This argument has been foreclosed by the law of the Circuit.

The “threshold question” of whether a plaintiff has Article III standing is whether the plaintiff has adequately “alleged an ‘injury-in-fact’ that is fairly traceable to the challenged conduct and redressable by a favorable judicial decision.” *Lerman v. Bd. of Elections*, 232 F.3d 135, 143 n.9 (2d Cir. 2000). This question must remain “distinct” from “the question of whether [plaintiffs have] valid claim on the merits.” *Id.*; see also *Carver v. City of New York*, 621 F.3d 221, 226 (2d Cir. 2010); *Denney v. Deutsche Bank AG*, 443 F.3d 253, 265 (2d Cir. 2006) (finding that class members “suffered injuries-in-fact[] irrespective of whether their injuries [were] sufficient to sustain any cause of action”). Defendants appear to conflate these two separate inquiries. See Defs. Decert. Mem. at 26–27. Before the Court is only the threshold question of standing.

In the class action context, the Supreme Court recently clarified that “[e]very class member must have Article III standing in order to *recover* individual damages.” *TransUnion v. Ramirez*, 141 S. Ct. 2190, 2208 (2021) (emphasis added). The Supreme Court’s holding did not address the “distinct question” of whether each member of the class must demonstrate standing

²³ Defendants also argue that Plaintiffs abandoned the *Int’l Bhd. of Teamsters v. United States*, 431 U.S. 324 (1977) framework when they did not pursue class certification for injunctive relief and that the Court relied on Plaintiffs seeking to certify an injunctive class to find standing in the Class Certification Order. Defs. Decert. Mem. at 25–27. Defendants are incorrect. In certifying the class, the Court did not rely on the existence of an injunctive class under Rule 23(b)(2). In fact, the Court explicitly stated “[c]lass claims for injunctive and declaratory relief under Rule 23(b)(2) and (c)(4) are now being pursued on a separate track—currently scheduled to be fully submitted in late 2018.” Class Cert. Order at 12–13. The Court declined to consider any arguments about a class under Rule 23(c)(4) because the Class Certification Order was limited to the Rule 23(b)(3) class. *Id.* at 13 n.3.

during class certification. *See id.* at 2208 n.4. However, the Second Circuit, has held that courts in this Circuit “do not require that each member of a class submit evidence of personal standing.” *Denney*, 443 F.3d at 263 (collecting cases). Instead, the class must “be defined in such a way that anyone within it would have standing.” *Id.* at 264.

This Circuit has not clearly articulated a test to assess standing for Title VII claims. However, in the equal protection context, the Supreme Court has held that to establish standing, a plaintiff must only show that a policy “prevented [the plaintiff] from competing on an equal footing,” not that the plaintiff “would have obtained a benefit but for the barrier.” *Ne. Fla. Chapter of Associated Gen. Contractors of Am. v. Jacksonville*, 508 U.S. 656, 666–67 (1993). The district court in *Houser v. Pritzker* held that this standard should also apply in the Title VII context as “federal courts generally are less reluctant to overcome jurisdictional hurdles when faced with statutory, rather than constitutional, merits questions.” 28 F. Supp. 3d 222, 238–39 (S.D.N.Y. 2014). The Court agrees and adopts the equal protection standing test for the Title VII claims in this case.

Combining the requirements for standing for classes and for Title VII claims, the test for standing in this case asks whether the class is defined in such a way as to ensure that the class members were subject to the policy that allegedly placed them on an unequal footing based on their gender. Because the class definition used here makes clear that all the class members were subject to the policies at issue, there is standing sufficient to maintain the certification of the class. *See Houser*, 28 F. Supp. 3d at 237.

2. Rule 23 Standard

Defendants renew their arguments made when certification of this case was initially briefed. Specifically, they allege that discovery has revealed facts that require the court to revisit

its decision as to commonality, predominance, typicality, superiority, and adequacy. Defs.

Decert. Mem. at 27–45. Again, the Court rejects Defendants’ arguments.²⁴

a. Commonality

Defendants contend that there are no common questions of fact because of the discretion involved in the compensation and promotion processes, *id.* at 28–37, and further argue that the class members work a number of different jobs, undermining commonality, *id.* at 31.

Defendants contend that discovery has only strengthened the arguments made at the class certification stage. *See id.* at 31. They adduce forty-one declarations from employees stating that discretion is exercised differently in implementing the processes. *Id.* at 4, 30. They also put forward a subset of those declarations and deposition excerpts to support their contention that the 360 review and quartiling processes are decentralized. *Id.* at 30. However, some of the declarations and depositions predate class certification. *See, e.g., id.* at vi–xiii; ECF Nos. 1225-1, 1225-2, 1225-4. Although Defendants argue that discovery has shown that discretion is exercised differently by different managers and evaluators, Defs. Decert. Mem. at 29–31, and there is no centralized decision-making, *id.* at 36, both arguments were raised—and rejected—when the Court certified the class, *see, e.g.* Defs. R&R Obj. at 1, 11–15, 22–25, 29–30, ECF No. 502. Plaintiffs counter that the processes constitute a common mode of decision-making as contemplated by *Dukes*, Pls. Decert. Opp’n at 29–31, and that the class members hold two positions, Associate and Vice President, *id.* at 31–32.

As the Court explained in the Class Certification Order, under *Dukes*, “‘a common mode of exercising discretion that pervades the entire company’ can constitute a general policy” for the

²⁴ Where Defendants make arguments in their decertification briefs that are more appropriate in relation to summary judgment than to decertification, the arguments are addressed in the summary judgment section of this order. *See* Defs. Decert. Reply at 19, ECF No. 1307 (arguing about liability for disparate treatment claims).

purposes of commonality. Class Cert. Order at 24 (quoting *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 356 (2011)). Plaintiffs' claims must "depend upon a common contention" that is "of such a nature that it is capable of classwide resolution—which means that determination of its truth or falsity will resolve an issue that is central to the validity of each one of the claims in one stroke." *Dukes*, 564 U.S. at 350. Further, the Supreme Court clarified that "giving discretion to lower-level supervisors can be the basis of Title VII liability under a disparate-impact theory" when that discretion is channeled through "specific employment practice[s]." *Id.* at 355, 357 (citation omitted).

First, Defendants argue that Plaintiffs must show that each class member was injured. *See* Defs. Decert. Mem. at 32–33. This argument misunderstands why the Court granted certification of the class. In Phase I of this case, the questions being asked are at the classwide level—did these processes cause disparate impact and does the use of the processes evidence disparate treatment? At Phase II, the case will look at individuals and whether they suffered harm such that damages should be awarded to them. All of the class members were subject to at least one of the processes. *See* Class Cert. Order at 22–49. The processes were a means of channeling the discretion of lower-level supervisors. *See Dukes*, 564 U.S at 355. All the class members have this in common and it forms the basis of their common contention. At Phase I, the common questions will "generate common *answers* apt to drive the resolution of the litigation." *Dukes*, 564 U.S. at 350 (quoting Richard A. Nagareda, *Class Certification in the Age of Aggregate Proof*, 84 N.Y.U. L. Rev. 97, 132 (2009)) (emphasis in original). In the Class Certification Order, the Court explained at length how Plaintiffs demonstrated commonality, *see*

Class Cert. Order at 23–34, and Defendants’ arguments for decertification do not raise issues that the Court finds persuasive or different from those previously considered.²⁵

Defendants also argue that the class members perform different jobs, Defs. Decert. Mem. at 31; Defs. Decert. Reply at 13–16, ECF No. 1307, an argument they made when class certification was before the Court, *see, e.g.*, Defs. R&R Obj. at 1–3, 12 n.5, 20, 25. The Court noted in its Class Certification Order that class members here “hold only one of two jobs: Associate or Vice President.” Class Cert. Order at 27–28 n.11. That continues to be true. Even if it is understood that there is some variability in the tasks each Associate or Vice President performs, each Associate and each Vice President is at the same level of “relative seniority” in the Goldman Sachs hierarchy. *See* Defs. 56.1 ¶¶ 7–8. This is different from the class members in *Dukes* who “held a multitude of different jobs, at different levels of Wal-Mart’s hierarchy” such that they had “little in common but their sex and this lawsuit.” *Dukes*, 564 U.S. at 359–60 (quoting *Dukes v. Wal-Mart Stores, Inc.*, 603 F.3d 571, 652 (9th Cir. 2010) (Kozinski, J., dissenting)). These factual distinctions are critical, and place this case outside *Dukes*’ ambit.

The Court finds that Plaintiffs, as they did in 2018, have met their burden of demonstrating commonality.

²⁵ For example, Defendants’ argument related to *Mandala* is unavailing. Defs. Decert. Mem. at 32–33 (citing *Mandala v. NTT Data, Inc.*, 975 F.3d 202 (2d Cir. 2020)). In *Mandala*, the plaintiffs used nationwide statistics that were not properly tailored to the inquiry. 975 F.3d at 210–11. By contrast, Plaintiffs in this case use data from employees at Goldman Sachs, which is properly tailored to the inquiry. *See generally* Farber Report. Defendants also cite the certification decision in *Kassman v. KPMG LLP*, 416 F. Supp. 3d 252 (S.D.N.Y. 2018), for the proposition that there is no common question of fact when “local managers apply highly discretionary criteria.” Defs. Decert. Mem. at 28; *see also id.* at 29–31. But, in *Kassman*, unlike here, the policies dictated “stages or phases” at which discretion was exercised rather than channeling the discretion. 416 F. Supp. 3d at 278. *Moussouris v. Microsoft*, No. 15 Civ. 1483, 2018 WL 3328418 (W.D. Wash. June 25, 2018), is also distinguishable from the case before the Court. In *Moussouris*, the plaintiffs argued that the root of the commonality was the unfettered discretion given to supervisors. 2018 WL 3328418, at *18. The court held this argument to be foreclosed by *Dukes*. *Id.* At oral argument in *Moussouris* the plaintiffs did an about-face and argued that there was no discretion, *id.* This is dissimilar from the arguments Plaintiffs make here consistently, that discretion was channeled through the challenged processes. The other cases relied on by Defendants predate the Class Certification Order, and were already considered by the Court.

b. Typicality

Defendants argue that the typicality requirement is not met because “no class members’ claims . . . arose from the ‘same course of events.’”’ Defs. Decert. Mem. at 40 (quoting *Moussouris*, 2018 WL 3328418, at *27–28). They argue, as they have previously, that the named Plaintiffs’ differing experiences at Goldman Sachs demonstrate that the claims of individual class members are not typical. *Id.* at 40–42. Defendants allege that discovery has “show[n] that no class members’ claims, including those of [n]amed Plaintiffs, arose from the ‘same course of events,’”’ *id.* at 40 (emphasis and citation omitted), however they do not point to any facts that were not known when the class was certified.

Typicality “is satisfied when the lead plaintiffs’ claims arise from the same series of events and find support in the same legal theories as the claims of all of the remaining class members.” *Houser*, 28 F. Supp. 3d at 245; *see also* Fed. R. Civ. P. 23(a)(3) (requiring that “the claims or defenses of the representative parties are typical of the claims or defenses of the class”). The named Plaintiffs, like the unnamed class members, were all subjected to at least one of the challenged processes and their claims are based on the same legal theories. *See* Class Cert. Order at 36–39. The Court rejected nearly identical typicality arguments made by Defendants in 2018, *id.* at 34–39, and there has been no “significant intervening event or a showing of compelling reasons to reexamine the question” the Court already answered, *Mazzei I*, 308 F.R.D. at 106 (quoting *Gulino II*, 907 F. Supp. 2d at 504). Again, the Court rejects Defendants’ arguments for the reasons previously stated.

c. Adequacy

Defendants argue that “the evidence now shows that this class action is infected with intra-class conflicts that demand decertification” because class members have evaluated other

class members as part of the challenged processes, destroying adequacy. Defs. Decert. Mem. at 45. This alleged intra-class conflict was knowable at class certification given how the processes function, for example, peers evaluated each other as part of the 360 review process, Defs. 56.1 ¶ 13. However, Defendants can now approximate how many class members provided feedback as part of the 360 review process and were managers or cross-ruffers. Defs. Decert. Mem. at 17, 45.

“In order to defeat a motion for certification, any conflicts between the class representative and members of the putative class must be ‘fundamental.’” *Ligon v. City of New York*, 288 F.R.D. 72, 80 (S.D.N.Y. 2013) (quoting *In re Flag Telecom Holdings, Ltd. Sec. Litig.*, 574 F.3d 29, 35 (2d Cir. 2009)).

The processes at issue channel the discretion of the individual evaluators and cross-ruffers, limiting their discretion. *See* Class Cert. Order at 25–28. The processes negate any possible fundamental conflict between the class representatives and class members because the allegations are not that any individual evaluators acted discriminatorily, but rather that the processes caused disparate impact or evinced disparate treatment. *Cf. In re Johnson*, 760 F.3d 66, 74 (D.C. Cir. 2014) (“If class members neutrally applied a flawed rating system and thereby reached a discriminatory result, then they were not themselves discriminating and therefore have no apparent interest that is in conflict with the attempt to prove other agents were denied promotions because of their race.”). Defendants have not shown that there is a fundamental conflict in the class such that adequacy is lacking.

d. Predominance

Defendants rely mainly on cases decided before the Class Certification Order to argue that determining who will be entitled to damages will predominate over classwide issues. *See*

Defs. Decert. Mem. at 37–40.²⁶ Defendants merely reiterate arguments made in 2018 when the Court certified the class, and point to no changes in the factual record that impacts predominance.

“Rule 23(b)(3) [] does *not* require a plaintiff seeking class certification to prove that each element of her claim is susceptible to classwide proof. What the rule does require is that common questions ‘*predominate* over any questions affecting only individual [class] members.’” *Amgen Inc. v. Conn. Ret. Plans & Tr. Funds*, 568 U.S. 455, 459 (2013) (quoting Fed. R. Civ. P. 23(b)(3); other quotation marks, citations, and alterations omitted). In the liability phase of this case, common questions predominate. The parties’ motions for summary judgment illustrate this predominance, as each party believes they are entitled to summary judgment, which would apply to the entire class. Issues relating to damages and the possible need for individualized findings are not relevant at this stage of the case where common questions of liability predominate.

e. Superiority

Defendants argue that a classwide finding of liability will not establish Defendants’ liability as to individual class members for purposes of damages calculations. Defs. Decert. Mem. at 42–44. They argue that this issue is exacerbated by Plaintiffs’ decision not to pursue an injunctive class, *id.* at 42, a decision that occurred after class certification, *id.* at 1. A finding of liability at the class level may streamline individual liability determinations. *See Robinson v. Metro-North Commuter R.R.*, 267 F.3d 147, 161–62 (2d Cir. 2001), *abrogation on other grounds*

²⁶ Defendants rely on *Olean Wholesale Grocery Coop. v. Bumble Bee Foods LLC*, 993 F.3d 774 (9th Cir. 2021). Defs. Decert. Mem. at 37. Any reliance on *Olean* is misplaced as the Ninth Circuit has vacated the opinion. *Olean Wholesale Grocery Coop. v. Bumble Bee Foods LLC*, 5 F.4th 950, 951 (9th Cir. 2021). The other case Defendants cite that was decided after class certification is an antitrust case. *In re Rail Freight Fuel Surcharge Antitrust Litig.*, 934 F.3d 619 (D.C. Cir. 2019). The predominance analysis in that case focused on whether each class member was injured. *Id.* at 625–26. As explained in the standing section of this opinion, the class definition is constructed in such a way to ensure each member suffered an injury-in-fact. *See supra* Part II.B.1.

recognized by Johnson v. Nextel Commc'ns Inc., 780 F.3d 128, 138 (2d Cir. 2015) (disparate impact); *Reynolds*, 685 F.3d at 203–04 (disparate treatment). As the Court explained in 2018, “a class action is superior to other available methods for fairly and efficiently adjudicating Plaintiffs’ . . . claims.” Class Cert. Order at 47–49 (quoting Fed. R. Civ. P. 23(b)(3)).

Defendants have not carried the “heavy burden to prove the necessity of either the drastic step of decertification or the less draconian but still serious step of limiting the scope of the class.” *Gulino II*, 907 F. Supp. 2d at 504 (quoting *Gordon*, 117 F.R.D. at 61). Accordingly, Defendants’ motion for decertification is DENIED.

III. Summary Judgment

Plaintiffs move for summary judgment on their *prima facie* case for some of their disparate impact claims. Pls. Summ. J. Mot., ECF No. 1247. Defendants move for summary judgment on Plaintiffs’ disparate impact and disparate treatment claims. Defs. Summ. J. Mot., ECF No. 1239.

A. Legal Standard

1. Summary Judgment

“The court shall grant summary judgment if the movant shows that there is no genuine dispute as to any material fact and the movant is entitled to judgment as a matter of law.” Fed. R. Civ. P. 56(a). The moving party bears the initial burden of pointing to evidence in the record, “including depositions, documents[,] . . . [and] affidavits or declarations,” Fed. R. Civ. P. 56(c)(1)(A), “which it believes demonstrate[s] the absence of a genuine issue of material fact,” *Celotex Corp. v. Catrett*, 477 U.S. 317, 323 (1986). The moving party may support an assertion that there is no genuine dispute by “showing . . . that [the] adverse party cannot produce admissible evidence to support the fact.” Fed. R. Civ. P. 56(c)(1)(B). The non-movant cannot

avoid summary judgment “through mere speculation or conjecture” or “by vaguely asserting the existence of some unspecified disputed material facts.” *Western World Ins. Co. v. Stack Oil, Inc.*, 922 F.2d 118, 121 (2d Cir. 1990) (internal quotation marks and citations omitted). On summary judgment, the “[C]ourt’s function is not to weigh the evidence, make credibility determinations or resolve issues of fact, but rather to determine whether, drawing all reasonable inferences from the evidence presented in favor of the non-moving party, a fair-minded jury could find in the non-moving party’s favor.” *Urbont v. Sony Music Ent.*, 831 F.3d 80, 88 (2d Cir. 2016) (quoting *Beatie v. City of New York*, 123 F.3d 707, 710–11 (2d Cir. 1997)).

“The same standard of review applies when the court is faced with cross-motions for summary judgment.” *Clear Channel Outdoor, Inc. v. City of New York*, 608 F. Supp. 2d 477, 492 (S.D.N.Y. 2009) (citing *Morales v. Quintel Entm’t, Inc.*, 249 F.3d 115, 121 (2d Cir. 2001)). In evaluating cross-motions for summary judgment, “[e]ach party’s motion must be reviewed on its own merits, and the Court must draw all reasonable inferences against the party whose motion is under consideration.” *Id.* (citing *Morales*, 249 F.3d at 121). However, “even when both parties move for summary judgment, asserting the absence of any genuine issues of material fact, a court need not enter judgment for either party.” *Morales*, 249 F.3d at 121 (citing *Heublein, Inc. v. United States*, 996 F.2d 1455, 1461 (2d Cir. 1993)).

Competing expert testimony does not preclude summary judgment, so long as the district court excludes one parties’ expert under *Daubert*. *Raskin*, 125 F.3d at 66–67. Alternatively, if both experts support the findings required for a *prima facie* case, there is no material dispute of fact and summary judgment may be granted. See *City of New York I*, 637 F. Supp. 2d at 94.

2. Title VII

a. Disparate Impact

For a Title VII disparate impact claim, the plaintiffs must prove, by a preponderance of the evidence, that the challenged practices have caused a disparate impact on the basis of sex. This is completed in a three-part analysis involving shifting evidentiary burdens. *Gulino v. New York State Educ. Dep’t.* (*Gulino I*), 460 F.3d 361, 382 (2d Cir. 2006). First, plaintiffs must establish a *prima facie* case. See 42 U.S.C. § 2000e–2(k). Plaintiffs must identify the employment practice or practices they allege have harmed them. *Id.* § 2000e–2(k)(1)(B)(i). They must then prove that the practice or practices were responsible for the alleged disparity. *Id.* Statistical proof is central to the showing of a *prima facie* case of disparate impact because disparate impact claims require that an employment practice “had a disparate effect on the protected group” regardless of intent. *Robinson*, 267 F.3d at 160 (2d Cir. 2001). Regression modeling is normally used because this statistical technique attempts to eliminate other possible causes, leaving only the studied event as the possible cause of the difference observed. *Ottaviani*, 875 F.2d at 366–67, 371. In employment discrimination cases, statistical significance at the 5% level, or approximately two standard deviations, is sufficient. *Smith*, 196 F.3d at 366; see also *Ottaviani*, 875 F.2d at 371. This means there is a 95% probability that the difference found is due to the variable of interest, rather than chance. *Ottaviani*, 875 F.2d at 371. The analysis should be done on the relevant level at which the process at issue was implemented, unless plaintiffs can explain why it is proper to exclude a portion of the employees subject to the policy. See *Smith*, 196 F.3d at 368–70.

Once plaintiffs establish a *prima facie* case, the burden shifts to defendants to rebut the *prima facie* case by attacking plaintiffs’ statistical proof, *Gulino I*, 460 F.3d at 382, or by

“demonstrat[ing] that the challenged practice is job related for the position in question and consistent with business necessity,” *id.* (citing 42 U.S.C. § 2000e-2(k)(1)(A)(i)). If defendants attack plaintiffs’ statistical proof, plaintiffs are free to challenge defendants’ rebuttal just as in any other lawsuit, including by “impeach[ing] the reliability of the statistical evidence, . . . offer[ing] rebutting evidence, or . . . disparag[ing] in arguments or in briefs the probative weight which the . . . evidence should be accorded.” *Cf. Watson v. Fort Worth Bank & Trust*, 487 U.S. 977, 996 (1988) (O’Connor, J.).

Finally, if defendants rebut the *prima facie* case using the business necessity option, the burden shifts back to plaintiffs to show that there are less discriminatory alternatives. *See Gulino I*, 460 F.3d at 382.

b. Disparate Treatment

Like disparate impact claims, pattern-or-practice disparate treatment claims are evaluated using a burden-shifting framework. *Int’l. Bhd. of Teamsters*, 431 U.S. at 360–62; *City of New York III*, 717 F.3d at 83. Plaintiffs make out a *prima facie* case if they can show that (1) “discrimination was the company’s standard operating procedure[,] the regular rather than the unusual practice,” and (2) “the discrimination was directed at a class of victims.” *City of New York III*, 717 F.3d at 83 (internal quotation marks omitted) (alterations in original). For this showing, “instances of discrimination against particular employees are relevant to show a policy of intentional discrimination, [but] they are not required; a statistical showing of disparate impact might suffice.” *Id.* at 84. “The statistical disparities supporting [a] . . . finding . . . [of] disparate impact [can] also serve[] to establish a *prima facie* case on the . . . claim of a pervasive pattern of discriminatory treatment.” *See id.* at 88 (emphasis added). Further, the statistical significance level is the same for disparate impact and disparate treatment claims—two standard deviations.

Waisome v. Port Auth. of N.Y. & N.J., 948 F.2d 1370, 1376 (2d Cir. 1991).

After plaintiffs establish a *prima facie* case, the burden “shifts to the employer ‘to rebut the presumption of discrimination.’” *City of New York III*, 717 F.3d at 84 (quoting *Texas Dep’t of Cnty. Affs. v. Burdine*, 450 U.S. 248, 254 (1981)). “The employer need only ‘articulate some legitimate, nondiscriminatory reason for’” its presumptively discriminatory actions. *Id.* (quoting *Burdine*, 450 U.S. at 253) (emphasis omitted). As the Second Circuit has explained, defendants may rebut the presumption by attacking the accuracy or adequacy of plaintiffs’ statistics, or, alternatively, “by accepting [] plaintiff[s]’ statistics and producing non-statistical evidence to show that it lacked such an intent.” *Id.* at 85.

If defendants do not successfully rebut plaintiffs’ *prima facie* case, “the presumption arising from an unrebutted *prima facie* case entitles the plaintiff[s] to prevail on the issue of liability and proceed directly to the issue of appropriate relief.” *Id.* at 87. If defendants are successful in their rebuttal, the presumption in favor of plaintiffs “drops out” and the case proceeds to trial. *Id.*; *St. Mary’s Honor Center v. Hicks*, 509 U.S. 502, 509–11 (1993). At trial, plaintiffs are afforded the “opportunity to prove by a preponderance of the evidence that the legitimate reasons offered by the defendant were not true reasons, but were a pretext for discrimination.” *Reeves v. Sanderson Plumbing Prod., Inc.*, 530 U.S. 133, 143 (2000).

3. NYCHRL

NYCHRL discrimination claims are analyzed under the same burden-shifting standard used for Title VII claims. *Farmer v. Shake Shack Enter.*, 473 F. Supp. 3d 309, 323–24 (S.D.N.Y. 2020). The NYCHRL, however, must be construed “broadly in favor of discrimination plaintiffs, to the extent that such a construction is reasonably possible.” *Albunio v. City of New York*, 947 N.E.2d 135, 137 (N.Y. 2011). Accordingly, claims brought under the NYCHRL are analyzed “independently from and more liberally than federal and state

counterparts.” *Loeffler v. Staten Island Univ. Hosp.*, 582 F.3d 268, 278 (2d Cir. 2009) (internal quotation marks and citation omitted). Interpretations of analogous state and federal statutes may “be used to aid in interpretation of [the NYCHRL]” to the extent such interpretations serve as “a floor below which” the NYCHRL cannot fall. *Id.* (citation omitted) (emphasis in original).

B. Plaintiffs’ Motion for Summary Judgment

Plaintiffs move for partial summary judgment on their disparate impact claims under Title VII and the NYCHRL, contending that they met their burden of making out a *prima facie* case. Pls. Summ. J. Mot. With respect to the 360 review and quartiling processes, they move for summary judgment from the beginning of the class period to the end of 2015, Pls. Summ. J. Mem. at 1, when Goldman Sachs stopped reporting numeric scores for 360 reviews, Farber Report ¶ 85, 109. For cross-ruffing, they move for summary judgment from the beginning of the class period to the present. *Id.*

Plaintiffs argue they have established a *prima facie* case on their disparate impact claim, which Defendants have failed to rebut. Specifically, Plaintiffs contend that even if Shaw’s expert opinion is not excluded, her testimony is insufficient to rebut their *prima facie* case, because she “concedes that there is a statistically significant gender gap” caused by the processes at issue. *Id.* at 2–3, 12. Plaintiffs base this assertion on Shaw’s Exhibit 68, *id.* at 12, which recreates Farber’s aggregated analysis and then adds more variables Shaw believes should be included. Shaw Report at App. J. The fact that Shaw is able to reproduce Farber’s modeling does not mean their reports are in agreement, nor does it mean that Shaw’s report is insufficient to rebut Plaintiffs’ *prima facie* case at this stage. Further, Shaw’s final analysis in Exhibit 68, which includes all the variables she believes should be included, finds statistically significant gender differences only for Vice Presidents in two divisions, *see id.*, which does not support

Plaintiffs' motion for summary judgment. Contrary to Plaintiffs' argument, Shaw and Farber present different regression models, which include different variables, resulting in different conclusions about disparate impact. *Compare* Shaw Report ¶ 29, with Farber Report ¶¶ 8, 10–14, 17–18. Shaw and Farber's reports create a material dispute of fact as to how the processes at issue impacted the class members.

Where, as here, two experts have put forth admissible opinions that disagree, summary judgment is improper. *Raskin*, 125 F.3d at 66–67. Because the “two contradictory expert opinions meet the requisite threshold of reliability, it is the function of the jury, utilizing the ‘conventional devices’ of ‘cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof,’ to determine which is the more trustworthy and credible.” *In re Zyprexa*, 489 F. Supp. 2d at 285 (quoting *Daubert*, 509 U.S. at 596); *see also Washington v. Kellwood Co.*, 105 F. Supp. 3d 293, 315 (S.D.N.Y. 2015) (quoting *Amorgianos v. Nat'l. R.R. Passenger Corp.*, 303 F.3d 256, 267 (2d Cir. 2002)). Accordingly, Plaintiffs' motion for summary judgment is DENIED.

C. Defendants' Motion

Defendants move for summary judgment on all of Plaintiffs' claims. Defs. Summ. J. Mot. They group their motion in three parts—disparate impact claims, disparate treatment claims, and issues relating to certain claims and class members. *See generally* Defs. Summ. J. Mem.

1. Disparate Impact Claims

Defendants argue that Plaintiffs have improperly grouped employment practices in a manner foreclosed by 42 U.S.C. § 2000e-2(k)(1)(B)(i). *See id.* at 20–23. They also contend that Plaintiffs' disparate impact claims fail because of gaps in the statistical evidence. *Id.* at 23–27.

Finally, Defendants argue that Plaintiffs have not shown that there is a less discriminatory alternative to the processes at issue, dooming their case. *Id.* at 27–31. The Court addresses each argument separately.

a. Identification of Employment Practices

Defendants argue that they are entitled to summary judgment because Plaintiffs failed to break down the 360 review and cross-ruffing processes into smaller subcomponents. *Id.* at 20–23. The Court disagrees.

Under Title VII, plaintiffs must identify the “particular employment practice that causes a disparate impact on the basis of . . . sex.” *Ricci v. DeStefano*, 557 U.S. 557, 578 (2009). Plaintiffs must further show that the elements of the decision-making process “are not capable of separate analysis.” 42 U.S.C. § 2000e-2(k)(1)(B)(i). “Whether a particular decision[-]making process is capable of separation for analysis largely turns on the details of the specific process and its implementation in a given case.” *Chin*, 685 F.3d at 154.

Plaintiffs have met their burden by identifying three processes in Defendants’ overall evaluation, compensation, and promotion system—360 review, quartiling, and cross-ruffing. *See, e.g.*, Pls. Summ. J. Mem. at 1. The three processes identified are not capable of further disaggregation, as the elements of each subcomponent are intertwined and are meaningless when separated.

The 360 review process involves multiple subcomponents, but each subcomponent works with the others to create an actionable employment process—Defendants do not utilize the subcomponents of the 360 review process independently to make employment decisions. Rather, it is only when the subcomponents are brought together does the 360 review process become an

“element of a respondent’s decisionmaking process.” 42 U.S.C. § 2000e-2(k)(1)(B)(i). Before the elements come together, they are just data points rather than an employment practice.

In the cross-ruffing process, the two subcomponents of the process work in tandem to determine if an individual will be promoted to Managing Director. As the Court has explained above, *supra* Part I.B.3.b, the two subcomponents are considered one process for the purpose of 42 U.S.C. § 2000e-2(k)(1)(B)(i). *See Chin*, 685 F.3d at 154–55.

What Defendants ask the Court to do is akin to requiring that plaintiffs identify which question on an exam is discriminatory rather than making a claim as to the full exam. Defendants cite no precedent requiring such disaggregation. *See Stagi v. Nat'l R.R. Passenger Corp.*, 391 F. App'x. 133, 146–48 (3d Cir. 2010) (noting “[t]here is no legal requirement to use the smallest possible unit of analysis” in disparate impact cases). Breaking down the processes into their subcomponents would also serve to disaggregate the data, making recognizing statistically significant patterns more difficult. *See Ellis v. Costco Wholesale Corp.*, 285 F.R.D. 492, 523 n.25 (N.D. Cal. 2012) (“Indeed, Defendant appears to largely offer a tautology—that disaggregated data is more probative simply because it yields results favorable to [it].”). Plaintiffs have adequately defined the elements of the decision-making process such that they are not capable of meaningfully being further broken down. *See Chin*, 685 F.3d at 154–55.

b. Statistical Causation

Defendants allege that Plaintiffs have not met their burden of proving the employment processes at issue caused the disparities in compensation and promotions between men and women in the three divisions. Defs. Summ. J. Mem. at 23–27. Like Plaintiffs’ motion for summary judgment, Defendants’ motion must be denied because there is a disagreement between the experts on a number of material factual issues, including the proper level of aggregation and

how to analyze the cross-ruffing process, creating a dispute that must be resolved by a jury. *In re Zyprexa*, 489 F. Supp. 2d at 285. Defendants also argue that there is no material dispute of fact because Farber did not find proof of disparate impact after 2015 for the 360 review and quartiling processes and did not find that fewer women than men were promoted to Managing Director in every year. Defs. Summ. J. Reply at 7–8, ECF No. 1319. However, the change in impact of the processes after 2015 does not negate the statistical findings from before that year, especially given that Defendants adduced evidence that there were changes to the 360 review and quartiling processes starting in 2016. Landman Decl. ¶¶ 6, 20; Landman Dep. at 82, 88. In 2016, the 360 review process shifted to being primarily used as a development tool rather than used to determine an employee’s compensation. Landman Decl. ¶¶ 6, 20, 26–28; Landman Dep. at 82, 88. Goldman Sachs also changed its quartiling process in 2016. See Landman Decl. ¶¶ 26–28. Therefore, the lack of statistically significant findings after 2015 does not preclude a finding of liability prior to 2016. As to cross-ruffing, given the small sample size, there is a dispute among the experts about whether the statistical analysis should be done by year or across the full period. See Farber Rebuttal ¶ 81. Given the material disputes of fact, summary judgment is improper. See *Raskin*, 125 F.3d at 66.

c. Less Discriminatory Alternative

Defendants next argue that summary judgment is appropriate because Plaintiffs have failed to raise a triable issue of fact on the question of a less discriminatory alternative. Defs. Summ. J. Mem. at 27–31. In doing so, Defendants move for summary judgment on the final step of the disparate impact burden-shifting framework, a step only reached “if the defendant meets the burden of showing the challenged practice is job related.” *Gulino I*, 460 F.3d at 382 (emphasis added). As discussed above, if Plaintiffs establish a *prima facie* case, then Defendants

may show there is a business necessity for the processes at issue. *Id.* Only after this showing is made are Plaintiffs required to put forward a less discriminatory alternative. *Id.* Defendants have not defined the business necessity of the processes. Thus, even assuming, *arguendo*, that Plaintiffs have not sufficiently proposed a less discriminatory alternative, they cannot be faulted for that omission at this stage of the litigation. Although parties may move for summary judgment as to a “part of [a] claim or defense,” Fed. R. Civ. P. 56(a), Defendants point to no authority from this Circuit allowing parties to skip to the final step of a burden-shifting framework without first carrying their burden in an intervening step.²⁷ The Court declines to allow Defendants to turn the burden-shifting framework on its head.

Accordingly, Defendants’ motion for summary judgment on Plaintiffs’ disparate impact claims are DENIED.

2. Disparate Treatment Claims

Defendants move for summary judgment on Plaintiffs’ disparate treatment claims, arguing that the statistical evidence cannot support an inference that discrimination was the standard operating procedure at Goldman Sachs. Defs. Summ. J. Mem. at 31–36. For the reasons already stated, summary judgment can neither be granted based on Defendants’ disparate impact statistical causation argument, nor on their disparate treatment argument. There are

²⁷ Defendants’ cited cases do not compel a finding that they are entitled to summary judgment on the less discriminatory alternative issue without first reaching that step in the burden shifting framework. Defs. Summ. J. Mem. at 30. In *Hernandez*, the court addressed the less discriminatory alternative issue in dicta after determining the plaintiff had neither met the burden of identifying a hiring process nor shown that a disparity existed. *Hernandez v. Off. of Comm’r of Baseball*, No. 18 Civ. 9035, 2021 WL 1226499, at *11–12 (S.D.N.Y. Mar. 31, 2021). In *El*, the court granted summary judgment on a showing of business necessity because the defendant’s evidence was not rebutted by plaintiff—the court then addressed the alternative policy issue in dicta. *El v. Se. Pa. Transp. Auth.*, 479 F.3d 232, 237, 239–49 (3d Cir. 2007). In *Williams*, the court first found that the defendant met its burden of showing business necessity before addressing the less discriminatory alternative issue. *Williams v. Wells Fargo Bank, N.A.*, 901 F.3d 1036, 1040–42 (8th Cir. 2018). Finally, *Smith* arose under the Age Discrimination in Employment Act (“ADEA”). *Smith v. City of Des Moines*, 99 F.3d 1466, 1468–69 (8th Cir. 1996). The court addressed business necessity before assuming, but not holding, that the less discriminatory alternative step of the Title VII disparate impact framework applies to ADEA cases. *Id.* at 1473–74.

factual disputes on a number of material issues, including the proper level of aggregation and how to analyze the cross-ruffing process, creating a dispute that must be resolved by a jury. *In re Zyprexa*, 489 F. Supp. 2d at 285.

Accordingly, Defendants' motion for summary judgment on Plaintiffs' disparate treatment claims is DENIED.

3. Discrete Issues as to Certain Claims and Class Members

In addition to their arguments regarding Plaintiffs' disparate impact and treatment claims, Defendants move for summary judgment as to certain claims and class members. Defs. Summ. J. Mem. at 37–41. The Court shall address each of these miscellaneous arguments individually.

First, Defendants' argument that they are entitled to summary judgment as to cross-ruffing, *id.* at 37–38, arises from a misunderstanding of what the Court held in its Class Certification Order regarding promotions. Specifically, Defendants argue that Plaintiffs present no evidence related to “cross-ruffing” as Defendants define it. *Id.* As the Court has explained, cross-ruffing, for the purposes of this case, is the full Vice President-to-Managing Director promotion process, a process Farber did study, *see* Farber Report ¶¶ 125–35. Because Farber and Shaw disagree on whether the cross-ruffing process, as defined by the Court, caused a gender-based disparity, summary judgment is inappropriate. *See In re Zyprexa*, 489 F. Supp. 2d at 285 (quoting *Daubert*, 509 U.S. at 596).

Next, Defendants argue that summary judgment is warranted with respect to the 360 review process before 2005, because the process was not fully implemented until 2004 and there was insufficient data until 2005. Defs. Summ. J. Mem. at 38–39. Because the Court excluded Farber's testimony on the 360 review process for 2002 to 2004, Plaintiffs have not adduced any evidence of disparate impact or treatment due to this process for those three years. *See*

Robinson, 267 F.3d at 160 (disparate impact); *City of New York III*, 717 F.3d at 84 (disparate treatment). Accordingly, Defendants' motion for summary judgment for the 360 review process for those three years only is GRANTED.

Defendants also contend that they are entitled to summary judgment as to the 360 review and quartiling processes after 2015. Defs. Summ. J. Mem. at 39–40. The Court agrees. Farber has not put forward evidence of statistically significant differences based on gender after 2015. *See* Farber Report ¶ 114; *see also id.* ¶ 10. Plaintiffs have, therefore, failed to support their *prima facie* case. *See Robinson*, 267 F.3d at 160 (disparate impact); *City of New York III*, 717 F.3d at 84 (disparate treatment). The question of how the processes impacted pay after the processes were no longer in use because of historical pay differences carrying forward is an issue for damages in Phase II of this case. Accordingly, Defendants' motion for summary judgment for the 360 review and quartiling processes after 2015 only is GRANTED.

Defendants next contend that they are entitled to summary judgment for self-sustaining private wealth advisors, who work in the Investment Management division, because their compensation was not tied to the 360 review or quartiling processes. Defs. Summ. J. Mem. at 40–41. It is uncontested that self-sustaining private wealth advisors are not included in the normal compensation process because they are paid through commissions. Defs. 56.1 ¶ 70. However, female Associate and Vice President self-sustaining private wealth advisors were subjected to the 360 review and quartiling processes. Pls. 56.1 ¶ 1 (all Associates and Vice Presidents in the three divisions were subject to the policies). Although they may not recover damages given their differing compensation structure, they remain part of the class for the liability phase. Issues of individual damages are not currently before the Court, as they shall be addressed in Phase II of this case. Moreover, as to cross-ruffing, the Vice President self-

sustaining private wealth advisors are on equal footing as to promotions as are all other Vice Presidents in the class. Accordingly, Defendants' motion for summary judgment on this basis is DENIED.

Finally, Goldman Sachs moves for summary judgment as to three individuals who signed liability releases. As the individuals have released Defendants of liability as to claims "known or unknown," Defs. 56.1 ¶ 73, this motion is GRANTED.

Accordingly, Defendants' motion for summary judgment is GRANTED in part, and DENIED in part.

CONCLUSION

For the reasons stated above:

- Defendants' motion to preclude the testimony of Goldberg is DENIED.
- Defendants' motion to preclude the testimony of Butler as to paragraphs 62 through 67 of her report is GRANTED, and DENIED as to the remainder of her report.
- Defendants' motion to preclude the testimony of Farber as to the 360 review process for 2002 to 2004 and damages is GRANTED, and DENIED as to the remainder of his reports.
- Defendants' motion to preclude the testimony of Yermack is DENIED as moot.
- Plaintiffs' motion to preclude the testimony of Shaw as to disaggregation at the business unit level, the impact of removing outliers, and the subcomponents of the cross-ruffing process is GRANTED, and DENIED as to the remainder of her report.
- Plaintiffs' motion to preclude the testimony of Stockdale is DENIED.

- Plaintiffs' motion to preclude the testimony of Dunleavy as to the two statistical studies in his job relatedness reports and to prevent the Court from considering Dunleavy's job analysis report when assessing Defendants' decertification motion is DENIED.
- Plaintiffs' motion to preclude the testimony of Dunn is GRANTED.
- Defendants' motion for decertification is DENIED.
- Plaintiffs' motion for summary judgment is DENIED.
- Defendants' motion for summary judgment is GRANTED as to the 360 review process for 2002 to 2004, the 360 review and quartiling processes after 2015, and the three individuals who signed liability releases, and otherwise DENIED.

The Clerk of Court is directed to terminate the motions at ECF Nos. 1184, 1187, 1194, 1239, and 1247.

SO ORDERED.

Dated: March 17, 2022
New York, New York



ANALISA TORRES
United States District Judge